

TRAITE DE COOPERATION EN MATIERE DE BREVETS

PCT

NOTIFICATION D'ELECTION
(règle 61.2 du PCT)

Expéditeur: le BUREAU INTERNATIONAL

Destinataire:

United States Patent and Trademark
Office
(Box PCT)
Crystal Plaza 2
Washington, DC 20231
ÉTATS-UNIS D'AMÉRIQUE

en sa qualité d'office élu

Date d'expédition (jour/mois/année) 14 janvier 1999 (14.01.99)	
Demande internationale no PCT/FR98/01136	Référence du dossier du déposant ou du mandataire WOB97 CNR FLU
Date du dépôt international (jour/mois/année) 04 juin 1998 (04.06.98)	Date de priorité (jour/mois/année) 05 juin 1997 (05.06.97)
Déposant GALZI, Jean-Luc etc	

1. L'office désigné est avisé de son élection qui a été faite:

dans la demande d'examen préliminaire international présentée à l'administration chargée de l'examen préliminaire international le:

14 décembre 1998 (14.12.98)

dans une déclaration visant une élection ultérieure déposée auprès du Bureau international le:

2. L'élection a été faite

n'a pas été faite

avant l'expiration d'un délai de 19 mois à compter de la date de priorité ou, lorsque la règle 32 s'applique, dans le délai visé à la règle 32.b).

Bureau international de l'OMPI 34, chemin des Colombelettes 1211 Genève 20, Suisse no de télécopieur: (41-22) 740.14.35	Fonctionnaire autorisé Lazar Joseph Panakal no de téléphone: (41-22) 338.83.38
--	--

TRAITE DE COOPERATION EN MATIERE DE BREVETS

PCT

NOTIFICATION DE L'ENREGISTREMENT
D'UN CHANGEMENT(règle 92bis.1 et
instruction administrative 422 du PCT)

Date d'expédition (jour/mois/année)
04 mars 1999 (04.03.99)

Expéditeur: le BUREAU INTERNATIONAL

Destinataire:

GROSSET-FOURNIER, Chantal
Grosset-Fournier & Demachy S.A.R.L.
20, rue de Maubeuge
F-75009 Paris
FRANCE

Référence du dossier du déposant ou du mandataire
WOB97 CNR FLU

NOTIFICATION IMPORTANTE

Demande internationale no
PCT/FR98/01136

Date du dépôt international (jour/mois/année)
04 juin 1998 (04.06.98)

1. Les renseignements suivants étaient enregistrés en ce qui concerne:

le déposant l'inventeur le mandataire le représentant commun

Nom et adresse

GROSSET-FOURNIER, Chantal
Grosset-Fournier & Demachy S.A.R.L.
103, rue La Fayette
F-75481 Paris Cedex 10
FRANCE

Nationalité (nom de l'Etat)

Domicile (nom de l'Etat)

no de téléphone
331 42 81 09 58

no de télécopieur
331 42 81 08 71

no de téléimprimeur

2. Le Bureau international notifie au déposant que le changement indiqué ci-après a été enregistré en ce qui concerne:

la personne le nom l'adresse la nationalité le domicile

Nom et adresse

GROSSET-FOURNIER, Chantal
Grosset-Fournier & Demachy S.A.R.L.
20, rue de Maubeuge
F-75009 Paris
FRANCE

Nationalité (nom de l'Etat)

Domicile (nom de l'Etat)

no de téléphone
331 42 81 09 58

no de télécopieur
331 42 81 08 71

no de téléimprimeur

3. Observations complémentaires, le cas échéant:

4. Une copie de cette notification a été envoyée:

<input checked="" type="checkbox"/> à l'office récepteur	<input type="checkbox"/> aux offices désignés concernés
<input type="checkbox"/> à l'administration chargée de la recherche internationale	<input checked="" type="checkbox"/> aux offices élus concernés
<input checked="" type="checkbox"/> à l'administration chargée de l'examen préliminaire international	<input type="checkbox"/> autre destinataire:

Bureau international de l'OMPI
34, chemin des Colombettes
1211 Genève 20, Suisse

Fonctionnaire autorisé:

Marie-José Devillard

no de télécopieur (41-22) 740.14.35

no de téléphone (41-22) 338.83.38

Translation

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

50003

09/445205

Applicant's or agent's file reference	FOR FURTHER ACTION See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)	
International application No. PCT/FR98/01136	International filing date (day/month/year) 04 June 1998 (04.06.1998)	Priority date (day/month/year) 05 June 1997 (05.06.1997)
International Patent Classification (IPC) or national classification and IPC G01N 33/58		
Applicant CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.
2. This REPORT consists of a total of 8 sheets, including this cover sheet.

This report is also accompanied by ANNEXES, i.e., sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

These annexes consist of a total of _____ sheets.

3. This report contains indications relating to the following items:

- I Basis of the report
- II Priority
- III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- IV Lack of unity of invention
- V Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement
- VI Certain documents cited
- VII Certain defects in the international application
- VIII Certain observations on the international application

Date of submission of the demand 14 December 1998 (14.12.1998)	Date of completion of this report 24 August 1999 (24.08.1999)
Name and mailing address of the IPEA/EP European Patent Office D-80298 Munich, Germany Facsimile No. 49-89-2399-4465	Authorized officer Telephone No. 49-89-2399-0

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/FR98/01136

I. Basis of the report

1. This report has been drawn on the basis of (*Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.*):

- the international application as originally filed.
- the description, pages 1-60, as originally filed,
pages _____, filed with the demand,
pages _____, filed with the letter of _____,
pages _____, filed with the letter of _____.
- the claims, Nos. 1-34, as originally filed,
Nos. _____, as amended under Article 19,
Nos. _____, filed with the demand,
Nos. _____, filed with the letter of _____,
Nos. _____, filed with the letter of _____.
- the drawings, sheets/fig 1/22-22/22, as originally filed,
sheets/fig _____, filed with the demand,
sheets/fig _____, filed with the letter of _____,
sheets/fig _____, filed with the letter of _____.

2. The amendments have resulted in the cancellation of:

- the description, pages _____
 the claims, Nos. _____
 the drawings, sheets/fig _____

3. This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).

4. Additional observations, if necessary:

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/FR98/01136

III. Non-establishment of opinion with regard to novelty, inventive step and industrial applicability

The questions whether the claimed invention appears to be novel, to involve an inventive step (to be non obvious), or to be industrially applicable have not been examined in respect of:

- the entire international application.
- claims Nos. 1-34

because:

- the said international application, or the said claims Nos. _____ relate to the following subject matter which does not require an international preliminary examination (*specify*): _____

- the description, claims or drawings (*indicate particular elements below*) or said claims Nos. 1-34 are so unclear that no meaningful opinion could be formed (*specify*):

See the Supplemental Box.

- the claims, or said claims Nos. _____ are so inadequately supported by the description that no meaningful opinion could be formed.
- no international search report has been established for said claims Nos. _____.

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No.

PCT/FR98/01136

VI. Certain documents cited**1. Certain published documents (Rule 70.10)**

Application No. Patent No.	Publication date (day/month/year)	Filing date (day/month/year)	Priority date (valid claim) (day/month/year)

See the Supplemental Box.**2. Non-written disclosures (Rule 70.9)**

Kind of non-written disclosure	Date of non-written disclosure (day/month/year)	Date of written disclosure referring to non-written disclosure (day/month/year)

INTERNATIONAL PRELIMINARY EXAMINATION REPORTInternational application No.
PCT/FR 98/01136**I. Basis of the report**

1. This report has been drawn on the basis of (*Replacement sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.*):

Preliminary remarks

In accordance with Rule 13^{ter}.1(f), any sequence listing not contained in the international application as filed shall not form part of the international application.

INTERNATIONAL PRELIMINARY EXAMINATION REPORTInternational application No.
PCT/FR 98/01136**Supplemental Box**
(To be used when the space in any of the preceding boxes is not sufficient)

Continuation of: III.

The present international preliminary examination report does not contain an opinion concerning industrial applicability, novelty and inventive step for the following reasons:

The claims in the present application flagrantly violate the requirements of PCT Article 6, which stipulates that claims must be clear. This applies to the claims taken individually or as a whole (cf. PCT Guidelines C-III, 4.1).

Although claims 1 and 22 / 2 and 23 / 10 and 11 / 30 and 31 / 33 and 34 have been formulated as distinct independent claims, they seem to have the same subject matter and to differ from one another only by a variation in the definition of the subject matter for which protection is sought. Consequently, these claims are not concise. Moreover, taken as a whole, they lack clarity because, owing to the large number of independent claims, it is difficult, if not impossible, to determine the subject matter for which protection is sought, and the determination by a third party of the scope of protection sought necessitates excessive efforts.

Further, the claims contain many features that are optional but are nonetheless repeated in several claims (cf. the part concerning GFP, for example). This makes the claims obscure and difficult to read, to say the least, and prevents precise characterization of their scope (cf. PCT Guidelines C-III, 3.7).

The applicant has also made excessive use of functional

INTERNATIONAL PRELIMINARY EXAMINATION REPORTInternational application No.
PCT/FR 98/01136**Supplemental Box**

(To be used when the space in any of the preceding boxes is not sufficient)

Continuation of III.

features in the claims (e.g. claim 1, "molecule able to absorb the light emitted", see PCT Guidelines C-III, 2.1). The same observation applies to the definition of the invention by the result to be achieved (cf. PCT Guidelines C-III, 4.7), which is found in most of the claims in the present application (cf. claim 1, "variants derived from GFP, by addition,..., provided that these variants retain the property of fluorescence"). Finally, the essential features of the invention are not clearly defined in the claims and the description does not enable this aspect to be clarified more precisely (cf. PCT Guidelines C-III, 3.4).

The combination of these breaches of PCT Articles and Guidelines prevents the establishment of an opinion concerning novelty, inventive step and industrial applicability.

INTERNATIONAL PRELIMINARY EXAMINATION REPORTInternational application No.
PCT/FR 98/01136**Supplemental Box**
(To be used when the space in any of the preceding boxes is not sufficient)

Continuation of VI.

Certain published documents (Rule 70.10)

Application No. Patent No.	Publication date (day/month/year)	Filing date (day/month/year)	Priority date (validly claimed) (day/month/year)
-------------------------------	--------------------------------------	---------------------------------	--

WO 97/28261	7.8.1997	31.1.1997	31.1.1996
-------------	----------	-----------	-----------

The applicant's attention is drawn to the fact that the document mentioned above, which enjoys a priority date earlier than that of the present application, will be able to be used in the European regional phase in so far as its content is relevant.

This examination report assumes that all claims enjoy the right of priority as of the filing date of the priority document. Should this later prove not to be the case, the following document - Nature, vol. 388, 1997, p. 882-887, cited in the international search report - could become relevant.

INTERNATIONAL PRELIMINARY EXAMINATION REPORTInternational application No.
PCT/FR 98/01136**VII. Certain defects in the international application**

The following defects in the form or contents of the international application have been noted:

1. Contrary to the requirements of PCT Rule 5.1(a)(ii), the description does not indicate the relevant prior art contained in documents D1-D3 and does not cite these documents.
2. The reference to figure 12b on page 52 is not correct (cf. pages 43 and 52).
3. The application number cited on page 1, line 22, violates PCT Guidelines C-II, 4.17.

VIII. Certain observations on the international application

The following observations on the clarity of the claims, description, and drawings or on the question whether the claims are fully supported by the description, are made:

1. The claims in the present application are not clear (PCT Article 6). The fluorescent proteins that are the subject of the present invention are cnidarian proteins with parameters defined e.g. in claim 1 (more than $14000\text{ M}^{-1}.\text{cm}^{-1}$ and more than 0.38 fluorescence quantum yield). The protein BFP is, on the one hand, excluded from this definition (cf. description in the application, page 11, line 11) and, on the other hand, forms part of the invention (cf. description of the application, page 4, line 21 and page 33, line 8, and see PCT Guidelines C-III, 4.3).
2. PCT Guidelines C-III, 2.1 stipulates that the claims must be drafted in terms of the technical features of the invention. References to scientific publications are not technical features (e.g. claim 26). Moreover, in claims 8, 14 and 25, the use of the term particularly ("notamment") makes the references optional and said references have no limiting effect on said claims (PCT Article 6, Rule 6).
3. Claims 1-3, 5, 6 and 22-24 are not clear for the following reasons (PCT Article 6). The introductory part of said claims refers to fluorescent cnidarian proteins defined by their molecular extinction coefficient of more than $14000\text{ M}^{-1}.\text{cm}^{-1}$ and their fluorescence quantum yield of more than 0.38. Since the use of the term particularly has no limiting effect on said fluorescent proteins, it follows that

VIII Certain observations on the international application

said claims are not limited to GFP, contrary to what seems to be suggested in the part of said claims dealing with detection. The claims therefore contain a contradiction which prevents their scope from being determined.

4. The feature of claim 30 according to which cells or cell fragments comprising the gene coding for a fluorescent substance fused with a ligand gene are claimed is not mentioned in the description. Claim 30 is not, therefore, supported by the description, as required under PCT Article 6.
5. The use of parentheses and square brackets in the claims should be limited to cases where it is absolutely necessary (cf. PCT Guidelines C-III, 4.11). It is not clear whether these parentheses and square brackets have a limiting effect on the claims contained in them (PCT Article 6).
6. Claim 7 does not meet the requirements of PCT Article 6. A choice of two different types of fluorescent protein can be used in claim 7, either EGFP or ECFP. However, the introductory part of this claim does not mention the fluorescent protein ECFP.
7. Claims 15 and 16 are not clear (PCT Article 6). These two claims mention a marked substance. The claims do not, however, describe the marked substance but only the marker (Bodipy and coumarin).
8. Claim 32 is not clear (PCT Article 6). The claimed

VIII. Certain observations on the international application

kit apparently contains a target protein coupled with any fluorescent protein (cf. page 79, line 7). However, the fluorescent protein previously described in the claim is selected from among fluorescent cnidarian proteins having the indicated optical properties.

9. The terms particularly, etc..., such that ("tel/telle que") and possibly ("éventuellement") used in claims 1, 2, 4-8, 10, 11, 13, 14, 17-23, 25, 27, 28, 31 and 32 are vague and ambiguous, and give rise to doubt as to the meaning of the technical features to which they refer. The same observation applies to the term about ("environ"), used to define the molecular extinction coefficient of more than $14000\text{ M}^{-1}.\text{cm}^{-1}$ and the fluorescence quantum yield of the autofluorescent cnidarian proteins of more than 0.38 in claims 1, 2, 4, 6, 10, 11, 21-24, 27 and 29-32. The subject of said claims is not, therefore, clearly defined (PCT Article 6 and Guidelines C-III, 4.5a and 4.6).

REPUBLIQUE FRANÇAISE

26 MARS 1998 - L E

09/445,205

[EPO 97 A TT CNR FLU]

INSTITUT NATIONAL
de la
PROPRIETE INDUSTRIELLE

RAPPORT DE RECHERCHE
PRELIMINAIRE

établi sur la base des dernières revendications
déposées avant le commencement de la recherche

N° d'enregistrement
nationalFA 544374
FR 9706977

DOCUMENTS CONSIDERES COMME PERTINENTS		Revendications concernées de la demande examinée
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	
T	MIYAWAKI A ET AL: "Fluorescent indicators for Ca-2+ based on green fluorescent proteins and calmodulin." NATURE (LONDON) 388 (6645). 1997. 882-887. ISSN: 0028-0836, XP002058386 * page 886, colonne 2 * ---	1-23
Y	ROMOSER V A ET AL: "Detection in living cells of Ca-2+-dependent changes in the fluorescence emission of an indicator composed of two green fluorescent protein variants linked by a calmodulin - binding sequence: A new class of fluorescent indicators." JOURNAL OF BIOLOGICAL CHEMISTRY 272 (20). 1997. 13270-13274. ISSN: 0021-9258, XP002058387 * le document en entier *	1-23
Y	PETUSHKOV V N ET AL: "Direct measurement of excitation transfer in the protein complex of bacterial luciferase hydroxyflavin and the associated yellow fluorescence proteins from Vibrio fischeri Y1." BIOCHEMISTRY 35 (25). 1996. 8413-8418. ISSN: 0006-2960, XP002058388 * le document en entier *	1-23
Y	MITRA R D ET AL: "Fluorescence resonance energy transfer between blue-emitting and red-shifted excitation derivatives of the green fluorescent protein." GENE (AMSTERDAM) 173 (1). 1996. 13-17. ISSN: 0378-1119, XP004042847 * le document en entier *	1-23
1	Date d'achèvement de la recherche 12 mars 1998	Examinateur Wells, A
CATEGORIE DES DOCUMENTS CITES		T : théorie où principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant
X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : pertinent à l'encontre d'au moins une revendication ou arrière-plan technologique général O : divulgation non-écrite P : document intercalaire		

REPUBLIC FRANÇAISE

**INSTITUT NATIONAL
de la
PROPRIETE INDUSTRIELLE**

RAPPORT DE RECHERCHE PRELIMINAIRE

établi sur la base des dernières revendications déposées avant le commencement de la recherche

**N° d'enregistrement
national**

FA 544374
FR 9706977

CATEGORIE DES DOCUMENTS CITES

- X : particulièrement pertinent à lui seul
- Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie
- A : pertinent à l'encontre d'au moins une revendication ou arrière-plan technologique général
- O : divulgation non-écrite
- P : document interne

T : théorie ou principe à la base de l'invention
E : document de brevet bénéficiant d'une date antérieure
à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure.
D : cité dans la demande
L : cité pour d'autres raisons

& : membre de la même famille, document correspondant

**ANNEXE AU RAPPORT DE RECHERCHE PRELIMINAIRE
RELATIF A LA DEMANDE DE BREVET FRANÇAIS NO.**

FA 544374
FR 9706977

La présente annexe indique les membres de la famille de brevets relatifs aux documents brevets cités dans le rapport de recherche préliminaire visé ci-dessus.

Lesdits membres sont contenus au fichier informatique de l'Office européen des brevets à la date du

Les renseignements fournis sont donnés à titre indicatif et n'engagent pas la responsabilité de l'Office européen des brevets, ni de l'Administration française

12-03-1998

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO 9101305 A	07-02-91	AU 6054590 A CA 2064766 A EP 0484369 A JP 5501862 T US 5683888 A	22-02-91 23-01-91 13-05-92 08-04-93 04-11-97
WO 9641166 A	19-12-96	US 5661035 A AU 6264396 A	26-08-97 30-12-96
US 5342789 A	30-08-94	EP 0505479 A WO 9109312 A	30-09-92 27-06-91
EP 0552108 A	21-07-93	CA 2087413 A JP 6066802 A US 5631169 A	18-07-93 11-03-94 20-05-97
US 5439797 A	08-08-95	EP 0537270 A WO 9200388 A	21-04-93 09-01-92
US 5625048 A	29-04-97	AU 4155096 A EP 0804457 A WO 9623810 A	21-08-96 05-11-97 08-08-96

RAPPORT DE RECHERCHE INTERNATIONALE

Demande Internationale No
PCT/FR 98/01136

A. CLASSEMENT DE L'OBJET DE LA DEMANDE

CIB 6 G01N33/58 C12Q1/25 C12N5/10 C12Q1/37

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)

CIB 6 G01N C12Q C12N

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés)

C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie °	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
P, X	WO 97 28261 A (UNIV. CALIFORNIA; AURORA BIOSCIENCES CORP) 7 août 1997 voir le document en entier ---	1-34
P, X	MIYAWAKI A ET AL: "Fluorescent indicators for Ca-2+ based on green fluorescent proteins and calmodulin." NATURE (LONDON) 388 (6645). 1997. 882-887. ISSN: 0028-0836, XP002058386 voir page 886, colonne 2 --- -/-	1-34

Voir la suite du cadre C pour la fin de la liste des documents

Les documents de familles de brevets sont indiqués en annexe

° Catégories spéciales de documents cités:

- "A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent
- "E" document antérieur, mais publié à la date de dépôt international ou après cette date
- "L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)
- "O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens
- "P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

"T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention

"X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément

"Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier

"&" document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

30 décembre 1998

Date d'expédition du présent rapport de recherche internationale

08/01/1999

Nom et adresse postale de l'administration chargée de la recherche internationale
Office Européen des Brevets, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Fonctionnaire autorisé

Fletcher, A

RAPPORT DE RECHERCHE INTERNATIONALE

Demande Internationale No
PCT/FR 98/01136

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
Y	ROMOSER V A ET AL: "Detection in living cells of Ca-2+-dependent changes in the fluorescence emission of an indicator composed of two green fluorescent protein variants linked by a calmodulin - binding sequence: A new class of fluorescent indicators." JOURNAL OF BIOLOGICAL CHEMISTRY 272 (20). 1997. 13270-13274. ISSN: 0021-9258, XP002058387 voir le document en entier ---	1-34
Y	PETUSHKOV V N ET AL: "Direct measurement of excitation transfer in the protein complex of bacterial luciferase hydroxyflavin and the associated yellow fluorescence proteins from Vibrio fischeri Y1." BIOCHEMISTRY 35 (25). 1996. 8413-8418. ISSN: 0006-2960, XP002058388 voir le document en entier ---	1-34
Y	MITRA R D ET AL: "Fluorescence resonance energy transfer between blue-emitting and red-shifted excitation derivatives of the green fluorescent protein." GENE (AMSTERDAM) 173 (1). 1996. 13-17. ISSN: 0378-1119, XP004042847 voir le document en entier ---	1-34
X	WO 91 01305 A (UNIV WALES MEDICINE) 7 février 1991	1-4, 6, 10, 11, 17, 18, 32
Y	voir exemples 3, 7	5, 7-9, 13-16, 19-21, 29, 33
Y	WO 96 41166 A (UNIV CALIFORNIA ; TSIEN ROGER Y (US); GONZALEZ JESUS E III (US)) 19 décembre 1996 voir page 33, ligne 30 - ligne 35 ---	1-11, 13-21, 29, 32, 33
Y	US 5 342 789 A (CHICK WILLIAM L ET AL) 30 août 1994 voir le document en entier ---	1-11, 13-21, 29, 32, 33
Y	EP 0 552 108 A (UNIV MARYLAND) 21 juillet 1993 voir le document en entier ---	1-11, 13-21, 29, 32, 33
		-/-

RAPPORT DE RECHERCHE INTERNATIONALEDemande Internationale No
PCT/FR 98/01136**C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS**

Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
Y	US 5 439 797 A (TSIEN ROGER Y ET AL) 8 août 1995 voir le document en entier ----	1-11, 13-21, 29, 32, 33
A	US 5 625 048 A (TSIEN ROGER Y ET AL) 29 avril 1997 voir le document en entier -----	1

RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Demande Internationale No

PCT/FR 98/01136

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)			Date de publication
WO 9728261 A	07-08-1997	AU 2250497 A EP 0877805 A			22-08-1997 18-11-1998
WO 9101305 A	07-02-1991	AU 6054590 A CA 2064766 A EP 0484369 A JP 5501862 T US 5683888 A			22-02-1991 23-01-1991 13-05-1992 08-04-1993 04-11-1997
WO 9641166 A	19-12-1996	US 5661035 A AU 6264396 A CA 2223927 A EP 0834074 A			26-08-1997 30-12-1996 19-12-1996 08-04-1998
US 5342789 A	30-08-1994	EP 0505479 A WO 9109312 A			30-09-1992 27-06-1991
EP 0552108 A	21-07-1993	CA 2087413 A JP 6066802 A US 5631169 A			18-07-1993 11-03-1994 20-05-1997
US 5439797 A	08-08-1995	AT 170980 T DE 69130171 D EP 0537270 A WO 9200388 A			15-09-1998 15-10-1998 21-04-1993 09-01-1992
US 5625048 A	29-04-1997	AU 4155096 A EP 0804457 A JP 10509881 T WO 9623810 A US 5777079 A			21-08-1996 05-11-1997 29-09-1998 08-08-1996 07-07-1998

TRAITEMENT DE COOPERATION EN MATIERE DE BREVETS

PCT

RAPPORT DE RECHERCHE INTERNATIONALE

(article 18 et règles 43 et 44 du PCT)

Référence du dossier du déposant ou du mandataire WO97 CNR FLU	POUR SUITE voir la notification de transmission du rapport de recherche internationale (formulaire PCT/ISA/220) et, le cas échéant, le point 5 ci-après A DONNER	
Demande internationale n° PCT/FR 98/ 01136	Date du dépôt international (<i>jour/mois/année</i>) 04/06/1998	(Date de priorité (la plus ancienne) (<i>jour/mois/année</i>) 05/06/1997
Déposant CENTRE NATIONAL DE LA RECHERCHE SCIENT.et al.		

Le présent rapport de recherche internationale, établi par l'administration chargée de la recherche internationale, est transmis au déposant conformément à l'article 18. Une copie en est transmise au Bureau international.

Ce rapport de recherche internationale comprend 4 feilles.

Il est aussi accompagné d'une copie de chaque document relatif à l'état de la technique qui y est cité.

1. **Il a été estimé que certaines revendications ne pouvaient pas faire l'objet d'une recherche**(voir le cadre I).
2. **Il y a absence d'unité de l'invention**(voir le cadre II).
3. La demande internationale contient la divulgation d'un **listage de séquence de nucléotides ou d'acides aminés** et la recherche internationale a été effectuée sur la base du listage de séquence
 - déposé avec la demande internationale
 - fourni par le déposant séparément de la demande internationale
 - sans être accompagnée d'une déclaration selon laquelle il n'inclut pas d'éléments allant au-delà de la divulgation faite dans la demande internationale telle qu'elle a été déposée.
 - transcrit par l'administration
4. En ce qui concerne le titre, le texte est approuvé tel qu'il a été remis par le déposant.
 Le texte a été établi par l'administration et a la teneur suivante:
5. En ce qui concerne l'**abrégé**,
 - le texte est approuvé tel qu'il a été remis par le déposant
 - le texte (reproduit dans le cadre III) a été établi par l'administration conformément à la règle 38.2b). Le déposant peut présenter des observations à l'administration dans un délai d'un mois à compter de la date d'expédition du présent rapport de recherche internationale.
6. La figure **des dessins** à publier avec l'abrégé est la suivante:
 Figure n° — suggérée par le déposant.
 parce que le déposant n'a pas suggéré de figure.
 parce que cette figure caractérise mieux l'invention.
 - Aucune des figures n'est à publier.

RAPPORT DE RECHERCHE INTERNATIONALE

Demande Internationale No

PCT/FR 98/01136

A. CLASSEMENT DE L'OBJET DE LA DEMANDE

CIB 6 G01N33/58 C12Q1/25 C12N5/10 C12Q1/37

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)

CIB 6 G01N C12Q C12N

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés)

C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
P, X	WO 97 28261 A (UNIV. CALIFORNIA; AURORA BIOSCIENCES CORP) 7 août 1997 voir le document en entier ---	1-34
P, X	MIYAWAKI A ET AL: "Fluorescent indicators for Ca-2+ based on green fluorescent proteins and calmodulin." NATURE (LONDON) 388 (6645). 1997. 882-887. ISSN: 0028-0836, XP002058386 voir page 886, colonne 2 ---	1-34



Voir la suite du cadre C pour la fin de la liste des documents



Les documents de familles de brevets sont indiqués en annexe

° Catégories spéciales de documents cités:

"A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent

"E" document antérieur, mais publié à la date de dépôt international ou après cette date

"L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)

"O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens

"P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

"T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention

"X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément

"Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier

"&" document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

30 décembre 1998

Date d'expédition du présent rapport de recherche internationale

08/01/1999

Nom et adresse postale de l'administration chargée de la recherche internationale

Office Européen des Brevets, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Fonctionnaire autorisé

Fletcher, A

RAPPORT DE RECHERCHE INTERNATIONALE

Demande Internationale No

PCT/FR 98/01136

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
Y	ROMOSER V A ET AL: "Detection in living cells of Ca-2+-dependent changes in the fluorescence emission of an indicator composed of two green fluorescent protein variants linked by a calmodulin - binding sequence: A new class of fluorescent indicators." JOURNAL OF BIOLOGICAL CHEMISTRY 272 (20). 1997. 13270-13274. ISSN: 0021-9258, XP002058387 voir le document en entier ---	1-34
Y	PETUSHKOV V N ET AL: "Direct measurement of excitation transfer in the protein complex of bacterial luciferase hydroxyflavin and the associated yellow fluorescence proteins from Vibrio fischeri Y1." BIOCHEMISTRY 35 (25). 1996. 8413-8418. ISSN: 0006-2960, XP002058388 voir le document en entier ---	1-34
Y	MITRA R D ET AL: "Fluorescence resonance energy transfer between blue-emitting and red-shifted excitation derivatives of the green fluorescent protein." GENE (AMSTERDAM) 173 (1). 1996. 13-17. ISSN: 0378-1119, XP004042847 voir le document en entier ---	1-34
X	WO 91 01305 A (UNIV WALES MEDICINE) 7 février 1991 ---	1-4, 6, 10, 11, 17, 18, 32
Y	voir exemples 3, 7 ---	5, 7-9, 13-16, 19-21, 29, 33
Y	WO 96 41166 A (UNIV CALIFORNIA ; TSIEN ROGER Y (US); GONZALEZ JESUS E III (US)) 19 décembre 1996 voir page 33, ligne 30 - ligne 35 ---	1-11, 13-21, 29, 32, 33
Y	US 5 342 789 A (CHICK WILLIAM L ET AL) 30 août 1994 voir le document en entier ---	1-11, 13-21, 29, 32, 33
Y	EP 0 552 108 A (UNIV MARYLAND) 21 juillet 1993 voir le document en entier ---	1-11, 13-21, 29, 32, 33
		-/-

RAPPORT DE RECHERCHE INTERNATIONALE

Demande Internationale No

PCT/FR 98/01136

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
Y	US 5 439 797 A (TSIEN ROGER Y ET AL) 8 août 1995 voir le document en entier ---	1-11, 13-21, 29, 32, 33
A	US 5 625 048 A (TSIEN ROGER Y ET AL) 29 avril 1997 voir le document en entier -----	1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/FR 98/01136

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
WO 9728261	A	07-08-1997	AU EP	2250497 A 0877805 A	22-08-1997 18-11-1998
WO 9101305	A	07-02-1991	AU CA EP JP US	6054590 A 2064766 A 0484369 A 5501862 T 5683888 A	22-02-1991 23-01-1991 13-05-1992 08-04-1993 04-11-1997
WO 9641166	A	19-12-1996	US AU CA EP	5661035 A 6264396 A 2223927 A 0834074 A	26-08-1997 30-12-1996 19-12-1996 08-04-1998
US 5342789	A	30-08-1994	EP WO	0505479 A 9109312 A	30-09-1992 27-06-1991
EP 0552108	A	21-07-1993	CA JP US	2087413 A 6066802 A 5631169 A	18-07-1993 11-03-1994 20-05-1997
US 5439797	A	08-08-1995	AT DE EP WO	170980 T 69130171 D 0537270 A 9200388 A	15-09-1998 15-10-1998 21-04-1993 09-01-1992
US 5625048	A	29-04-1997	AU EP JP WO US	4155096 A 0804457 A 10509881 T 9623810 A 5777079 A	21-08-1996 05-11-1997 29-09-1998 08-08-1996 07-07-1998

BP

TRAITE DE COOPERATION EN MATIERE DE BREVETS

PCT

REC'D 27 AUG 1999

WIPO PCT

RAPPORT D'EXAMEN PRELIMINAIRE INTERNATIONAL

(article 36 et règle 70 du PCT)

Référence du dossier du déposant ou du mandataire ---	POUR SUITE A DONNER voir la notification de transmission du rapport d'examen préliminaire international (formulaire PCT/IPEA/416)	
Demande internationale n° PCT/FR98/01136	Date du dépôt international (<i>jour/mois/année</i>) 04/06/1998	Date de priorité (<i>jour/mois/année</i>) 05/06/1997
Classification internationale des brevets (CIB) ou à la fois classification nationale et CIB G01N33/58		
Déposant CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE et al		
<p>1. Le présent rapport d'examen préliminaire international, établi par l'administration chargée de l'examen préliminaire international, est transmis au déposant conformément à l'article 36.</p> <p>2. Ce RAPPORT comprend 8 feuilles, y compris la présente feuille de couverture.</p> <p><input type="checkbox"/> Il est accompagné d'ANNEXES, c'est-à-dire de feuilles de la description, des revendications ou des dessins qui ont été modifiées et qui servent de base au présent rapport ou de feuilles contenant des rectifications faites auprès de l'administration chargée de l'examen préliminaire international (voir la règle 70.16 et l'instruction 607 des Instructions administratives du PCT).</p> <p>Ces annexes comprennent feuilles.</p>		
<p>3. Le présent rapport contient des indications relatives aux points suivants:</p> <ul style="list-style-type: none"> I <input checked="" type="checkbox"/> Base du rapport II <input type="checkbox"/> Priorité III <input checked="" type="checkbox"/> Absence de formulation d'opinion quant à la nouveauté, l'activité inventive et la possibilité d'application industrielle IV <input type="checkbox"/> Absence d'unité de l'invention V <input type="checkbox"/> Déclaration motivée selon l'article 35(2) quant à la nouveauté, l'activité inventive et la possibilité d'application industrielle; citations et explications à l'appui de cette déclaration VI <input checked="" type="checkbox"/> Certains documents cités VII <input checked="" type="checkbox"/> Irrégularités dans la demande internationale VIII <input checked="" type="checkbox"/> Observations relatives à la demande internationale 		

Date de présentation de la demande d'examen préliminaire internationale 14/12/1998	Date d'achèvement du présent rapport 24.08.99
Nom et adresse postale de l'administration chargée de l'examen préliminaire international: Office européen des brevets D-80298 Munich Tél. (+49-89) 2399-0 Tx: 523656 epmu d Fax: (+49-89) 2399-4465	Fonctionnaire autorisé Renggli, J N° de téléphone (+49-89) 2399



**RAPPORT D'EXAMEN
PRELIMINAIRE INTERNATIONAL**

Demande internationale n° PCT/FR98/01136

I. Base du rapport

1. Ce rapport a été rédigé sur la base des éléments ci-après (*les feuilles de remplacement qui ont été remises à l'office récepteur en réponse à une invitation faite conformément à l'article 14 sont considérées, dans le présent rapport, comme "initiallement déposées" et ne sont pas jointes en annexe au rapport puisqu'elles ne contiennent pas de modifications.*) :

Description, pages:

1-60 version initiale

Revendications, N°:

1-34 version initiale

Dessins, feuilles:

1/22-22/22 version initiale

2. Les modifications ont entraîné l'annulation :

- de la description, pages :
- des revendications, n°s :
- des dessins, feuilles :

3. Le présent rapport a été formulé abstraction faite (de certaines) des modifications, qui ont été considérées comme allant au-delà de l'exposé de l'invention tel qu'il a été déposé, comme il est indiqué ci-après (règle 70.2(c)) :

4. Observations complémentaires, le cas échéant :

III. Absence de formulation d'opinion quant à la nouveauté, l'activité inventive et la possibilité d'application industrielle

La question de savoir si l'objet de l'invention revendiquée semble être nouveau, impliquer une activité inventive (ne pas être évident) ou être susceptible d'application industrielle n'a pas été examinée pour ce qui concerne :

- l'ensemble de la demande internationale.
- les revendications n°s 1-34.

**RAPPORT D'EXAMEN
PRELIMINAIRE INTERNATIONAL**

Demande internationale n° PCT/FR98/01136

parce que :

- la demande internationale, ou les revendications n°s en question, se rapportent à l'objet suivant, à l'égard duquel l'administration chargée de l'examen préliminaire international n'est pas tenue effectuer un examen préliminaire international (*préciser*) :
- la description, les revendications ou les dessins (*en indiquer les éléments ci-dessous*), ou les revendications n°s 1-34 en question ne sont pas clairs, de sorte qu'il n'est pas possible de formuler une opinion valable (*préciser*) :
voir feuille séparée
- les revendications, ou les revendications n°s en question, ne se fondent pas de façon adéquate sur la description, de sorte qu'il n'est pas possible de formuler une opinion valable.
- il n'a pas été établi de rapport de recherche internationale pour les revendications n°s en question.

VI. Certain documents cités

1. Certains documents publiés (règle 70.10)

et / ou

2. Divulgations non écrites (règle 70.9)

voir feuille séparée

VII. Irrégularités dans la demande internationale

Les irrégularités suivantes, concernant la forme ou le contenu de la demande internationale, ont été constatées :

voir feuille séparée

VIII. Observations relatives à la demande internationale

Les observations suivantes sont faites au sujet de la clarté des revendications, de la description et des dessins et de la question de savoir si les revendications se fondent entièrement sur la description :

voir feuille séparée

Remarques préliminaires:

En vertu de la règle 13ter.1 f), tout listage de séquences qui ne figure pas dans la demande internationale telle qu' elle a été déposée ne fait pas partie de la demande internationale.

Section III:

La présente rapport d'examen préliminaire international ne contient pas d'opinion concernant l'applicabilité industrielle, la nouveauté et l'inventivité pour les raisons suivantes:

Les revendications de la présente demande contreviennent de façon flagrante aux conditions énoncées dans l'article 6 PCT qui stipulent que les revendications doivent être claires. Ceci s'applique aux revendications prises individuellement ou dans leur ensemble (cf. Directives C-III, 4.1 PCT).

En effet, bien que les revendications 1 et 22 / 2 et 23 / 10 et 11 / 30 et 31 / 33 et 34 aient été rédigées sous forme de revendications indépendantes distinctes, il semble qu'elles aient le même objet et qu'elles ne diffèrent l'une de l'autre que par une variation dans la définition de l'objet pour lequel la protection est demandée. Par conséquent, ces revendications ne sont pas concises. De plus, prises dans leur ensemble, elles sont dénuées de clarté, car du fait de la pluralité des revendications indépendantes, il est difficile, voire impossible de déterminer l'objet pour lequel une protection est demandée, et la délimitation par un tiers de l'étendue de la protection demandée nécessite des efforts excessifs.

De plus, les revendications contiennent une multitude de caractéristiques optionnelles qui sont néanmoins répétées dans de nombreuses revendications (cf. partie concernant la GFP p. ex.). Ceci rend les revendications obscures, rend la lecture de celles-ci pour le moins ardue et ne permet pas de caractériser précisément leur étendue (cf. Directives C-III, 3.7 PCT).

Le demandeur a également fait recours de façon exagérée dans les revendications à des caractéristiques fonctionnelles (e.g. revendication 1,

molécule susceptible d'absorber la lumière émise, voir Directives C-III, 2.1 PCT). La même remarque s'applique à la définition de l'invention par le résultat recherché (cf. Directives C-III, 4.7 PCT) qui est retrouvé dans la plupart des revendications de la présente demande (cf. revendication 1, des variants dérivés de la GFP, par addition,..., sous réserve que ces variants conservent la propriété de fluorescence). Finalement, les caractéristiques essentielles à l'invention ne sont pas clairement définies dans les revendications et la description ne permet pas de clarifier ce point plus précisément (cf. Directives C-III, 3.4 PCT).

La combinaison de ces manquements aux Directives et Articles PCT ne permet pas l'établissement d'une opinion concernant la nouveauté, l'inventivité et l'applicabilité industrielle.

Section VI

Certains documents publiés (règle 70.10)

Demande n° Brevet n°	Date de publication (jour/mois/année)	Date de dépôt (jour/mois/année)	Date de priorité (valablement revendiquée) (jour/mois/année)
WO 97/28261	7.8.1997	31.1.1997	31.1.1996

L'attention du demandeur est attirée sur le fait que le document mentionné ci-dessus, au bénéfice d'une date de priorité antérieure à la présente demande, pourra être utilisé dans la phase régionale européenne pour autant que son contenu soit pertinent.

Cet examen est fondé sur l'hypothèse que toutes les revendications jouissent d'un droit de priorité datant du dépôt du document de priorité. S'il s'avérait qu'il n'en était pas ainsi, le document suivant, Nature, vol. 388, 1997, p. 882-887, cité dans le rapport de recherche internationale pourrait devenir pertinent.

Section VII

1. Contrairement à ce qu'exige la règle 5.1 a) ii) PCT, la description n'indique pas l'état de la technique antérieure pertinent exposé dans les documents D1-D3 et ne cite pas ces documents.
2. La référence à la figure 12b à la page 52 n'est pas correcte (cf. pages 43 et 52).
3. Le numéro d'application cité à la page 1, ligne 22 contrevient aux Directives C-II, 4.17 PCT.

Section VIII

1. Les revendications de la présente demande ne sont pas claires (Art. 6 PCT). Les protéines fluorescentes objet de la présente invention sont les protéines de cnidaires présentant les paramètres définis e.g. dans la revendication 1 (plus de $14000\text{ M}^{-1}.\text{cm}^{-1}$ et plus de 0,38 de rendement quantique de fluorescence). La protéine BFP est, d'une part exclue de cette définition (cf. description de la demande, page 11, ligne 11) et d'autre part, fait partie de l'invention (cf. description de la demande, page 4, ligne 21 et page 33, ligne 8 et voir Directives C-III, 4.3 PCT).
2. Les directives PCT, C-III, 2.1 stipulent que les revendications doivent exposer les caractéristiques techniques de l'invention. Les références à des publications scientifiques ne sont pas des caractéristiques techniques (e.g. revendication 26). De plus, dans les revendications 8, 14 et 25 l'emploi du terme notamment rend les références optionnelles et celles-ci n'ont pas de portée limitative sur lesdites revendications (Art. 6, Règle 6 PCT).
3. Les revendications 1-3, 5, 6, 22-24 ne sont pas claires pour les raisons suivantes (Art. 6 PCT). La partie introductory desdites revendications fait référence à des protéines fluorescentes de cnidaires définies par leur coefficient d'extinction moléculaire supérieur à $14000\text{ M}^{-1}.\text{cm}^{-1}$ et le rendement quantique de fluorescence supérieur à 0.38. L'emploi du terme notamment n'ayant pas de

fonction limitative sur lesdites protéines fluorescentes, il s'ensuit que lesdites revendications **ne sont pas limitées** à la GFP, contrairement à ce qu'il semble être suggéré dans la partie consacrée à la détection des dites revendications. Ainsi, les revendications comportent une contradiction qui ne permet pas de déterminer leur étendue.

4. La caractéristique de la revendication 30, selon laquelle des cellules ou fragments de cellules comprenant le gène codant pour une substance fluorescente fusionnée avec le gène d'un ligand sont revendiquées, n'est pas mentionnée dans la description. La revendication 30 ne se fonde donc pas sur la description, comme l'exige l'article 6 PCT.
5. L'emploi de parenthèses et crochets dans les revendications doit être limité aux cas où cela est absolument nécessaire (cf. Directives, C-III, 4.11 PCT). Il n'est pas clair si ces parenthèses et crochets sont limitatifs pour les revendications en contenant (Article 6 PCT).
6. La revendication 7 ne satisfait pas aux conditions énoncées dans l'article 6 PCT. Deux types différents de protéine fluorescente sont utilisés à choix dans la revendication 7, soit EGFP ou ECFP. Cependant, la partie introductive de cette revendication ne fait pas mention de la protéine fluorescente ECFP.
7. Les revendications 15 et 16 ne sont pas claires (article 6 PCT). Ces deux revendications font mention d'une substance marquée. Ces revendications ne décrivent cependant pas la substance marquée mais le marqueur seul (Bodipy et coumarine).
8. La revendication 32 n'est pas claire (article 6 PCT). La trousse revendiquée contient apparemment une protéine cible couplée à une protéine fluorescente quelconque (cf. page 79, ligne 7). Cependant, la protéine fluorescente décrite auparavant dans la revendication est choisie parmi des protéines fluorescentes de cnidaires ayant les propriétés optiques indiquées.
9. Les termes notamment, etc..., tel/telle que, et éventuellement utilisés dans les revendications 1, 2, 4-8, 10, 11, 13, 14, 17-23, 25, 27, 28, 31 et 32 sont vagues et

équivoques, et laissent un doute quant à la signification des caractéristiques techniques auxquelles ils se réfèrent. La même remarque s'applique au terme environ employé pour définir le coefficient d'extinction moléculaire supérieur à $14000\text{ M}^{-1}.\text{cm}^{-1}$ et le rendement quantique de fluorescence supérieur à 0.38 des protéines autofluorescentes de cnidaires dans les revendications 1, 2, 4, 6, 10, 11, 21-24, 27, 29-32. L'objet desdites revendications n'est donc pas clairement défini (article 6 PCT et directives C-III, 4.5a et 4.6 PCT).

TRAITE DE COOPERATION EN MATIERE DE BREVETS

PCT

RAPPORT D'EXAMEN PRELIMINAIRE INTERNATIONAL

(article 36 et règle 70 du PCT)

Référence du dossier du déposant ou du mandataire ---	POUR SUITE A DONNER <small>voir la notification de transmission du rapport d'examen préliminaire international (formulaire PCT/IPEA/416)</small>	
Demande internationale n° PCT/FR98/01136	Date du dépôt international (<i>jour/mois/année</i>) 04/06/1998	Date de priorité (<i>jour/mois/année</i>) 05/06/1997
Classification internationale des brevets (CIB) ou à la fois classification nationale et CIB G01N33/58		
Déposant CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE et al		
<p>1. Le présent rapport d'examen préliminaire international, établi par l'administaration chargée de l'examen préliminaire international, est transmis au déposant conformément à l'article 36.</p> <p>2. Ce RAPPORT comprend 8 feuilles, y compris la présente feuille de couverture.</p> <p><input type="checkbox"/> Il est accompagné d'ANNEXES, c'est-à-dire de feuilles de la description, des revendications ou des dessins qui ont été modifiées et qui servent de base au présent rapport ou de feuilles contenant des rectifications faites auprès de l'administration chargée de l'examen préliminaire international (voir la règle 70.16 et l'instruction 607 des Instructions administratives du PCT).</p> <p>Ces annexes comprennent feuilles.</p>		
<p>3. Le présent rapport contient des indications relatives aux points suivants:</p> <ul style="list-style-type: none"> I <input checked="" type="checkbox"/> Base du rapport II <input type="checkbox"/> Priorité III <input checked="" type="checkbox"/> Absence de formulation d'opinion quant à la nouveauté, l'activité inventive et la possibilité d'application industrielle IV <input type="checkbox"/> Absence d'unité de l'invention V <input type="checkbox"/> Déclaration motivée selon l'article 35(2) quant à la nouveauté, l'activité inventive et la possibilité d'application industrielle; citations et explications à l'appui de cette déclaration VI <input checked="" type="checkbox"/> Certains documents cités VII <input checked="" type="checkbox"/> Irrégularités dans la demande internationale VIII <input checked="" type="checkbox"/> Observations relatives à la demande internationale 		

Date de présentation de la demande d'examen préliminaire internationale 14/12/1998	Date d'achèvement du présent rapport 24.08.99
Nom et adresse postale de l'administration chargée de l'examen préliminaire international:  Office européen des brevets D-80298 Munich Tél. (+49-89) 2399-0 Tx: 523656 epmu d Fax: (+49-89) 2399-4465	Fonctionnaire autorisé Renggli, J N° de téléphone (+49-89) 2399



**RAPPORT D'EXAMEN
PRELIMINAIRE INTERNATIONAL**

Demande internationale n° PCT/FR98/01136

I. Base du rapport

1. Ce rapport a été rédigé sur la base des éléments ci-après (*les feuilles de remplacement qui ont été remises à l'office récepteur en réponse à une invitation faite conformément à l'article 14 sont considérées, dans le présent rapport, comme "initiallement déposées" et ne sont pas jointes en annexe au rapport puisqu'elles ne contiennent pas de modifications.*) :

Description, pages:

1-60 version initiale

Revendications, N°:

1-34 version initiale

Dessins, feuilles:

1/22-22/22 version initiale

2. Les modifications ont entraîné l'annulation :

- de la description, pages :
 des revendications, n°s :
 des dessins, feuilles :

3. Le présent rapport a été formulé abstraction faite (de certaines) des modifications, qui ont été considérées comme allant au-delà de l'exposé de l'invention tel qu'il a été déposé, comme il est indiqué ci-après (règle 70.2(c)) :

4. Observations complémentaires, le cas échéant :

III. Absence de formulation d'opinion quant à la nouveauté, l'activité inventive et la possibilité d'application industrielle

La question de savoir si l'objet de l'invention revendiquée semble être nouveau, impliquer une activité inventive (ne pas être évident) ou être susceptible d'application industrielle n'a pas été examinée pour ce qui concerne :

- l'ensemble de la demande internationale.
 les revendications n°s 1-34.

**RAPPORT D'EXAMEN
PRELIMINAIRE INTERNATIONAL**

Demande internationale n° PCT/FR98/01136

parce que :

- la demande internationale, ou les revendications n°s en question, se rapportent à l'objet suivant, à l'égard duquel l'administration chargée de l'examen préliminaire international n'est pas tenue effectuer un examen préliminaire international (*PRECISER*) :
- la description, les revendications ou les dessins (*en indiquer les éléments ci-dessous*), ou les revendications n°s 1-34 en question ne sont pas clairs, de sorte qu'il n'est pas possible de formuler une opinion valable (*PRECISER*) :
voir feuille séparée
- les revendications, ou les revendications n°s en question, ne se fondent pas de façon adéquate sur la description, de sorte qu'il n'est pas possible de formuler une opinion valable.
- il n'a pas été établi de rapport de recherche internationale pour les revendications n°s en question.

VI. Certain documents cités

1. Certains documents publiés (règle 70.10)

et / ou

2. Divulgations non écrites (règle 70.9)

voir feuille séparée

VII. Irrégularités dans la demande internationale

Les irrégularités suivantes, concernant la forme ou le contenu de la demande internationale, ont été constatées :

voir feuille séparée

VIII. Observations relatives à la demande internationale

Les observations suivantes sont faites au sujet de la clarté des revendications, de la description et des dessins et de la question de savoir si les revendications se fondent entièrement sur la description :

voir feuille séparée

Remarques préliminaires:

En vertu de la règle 13ter.1 f), tout listage de séquences qui ne figure pas dans la demande internationale telle qu'elle a été déposée ne fait pas partie de la demande internationale.

Section III:

La présente rapport d'examen préliminaire international ne contient pas d'opinion concernant l'applicabilité industrielle, la nouveauté et l'inventivité pour les raisons suivantes:

Les revendications de la présente demande contreviennent de façon flagrante aux conditions énoncées dans l'article 6 PCT qui stipulent que les revendications doivent être claires. Ceci s'applique aux revendications prises individuellement ou dans leur ensemble (cf. Directives C-III, 4.1 PCT).

En effet, bien que les revendications 1 et 22 / 2 et 23 / 10 et 11 / 30 et 31 / 33 et 34 aient été rédigées sous forme de revendications indépendantes distinctes, il semble qu'elles aient le même objet et qu'elles ne diffèrent l'une de l'autre que par une variation dans la définition de l'objet pour lequel la protection est demandée. Par conséquent, ces revendications ne sont pas concises. De plus, prises dans leur ensemble, elles sont dénuées de clarté, car du fait de la pluralité des revendications indépendantes, il est difficile, voire impossible de déterminer l'objet pour lequel une protection est demandée, et la délimitation par un tiers de l'étendue de la protection demandée nécessite des efforts excessifs.

De plus, les revendications contiennent une multitude de caractéristiques optionnelles qui sont néanmoins répétées dans de nombreuses revendications (cf. partie concernant la GFP p. ex.). Ceci rend les revendications obscures, rend la lecture de celles-ci pour le moins ardue et ne permet pas de caractériser précisément leur étendue (cf. Directives C-III, 3.7 PCT).

Le demandeur a également fait recours de façon exagérée dans les revendications à des caractéristiques fonctionnelles (e.g. revendication 1,

molécule susceptible d'absorber la lumière émise, voir Directives C-III, 2.1 PCT). La même remarque s'applique à la définition de l'invention par le résultat recherché (cf. Directives C-III, 4.7 PCT) qui est retrouvé dans la plupart des revendications de la présente demande (cf. revendication 1, des variants dérivés de la GFP, par addition,..., sous réserve que ces variants conservent la propriété de fluorescence). Finalement, les caractéristiques essentielles à l'invention ne sont pas clairement définies dans les revendications et la description ne permet pas de clarifier ce point plus précisément (cf. Directives C-III, 3.4 PCT).

La combinaison de ces manquements aux Directives et Articles PCT ne permet pas l'établissement d'une opinion concernant la nouveauté, l'inventivité et l'applicabilité industrielle.

Section VI

Certains documents publiés (règle 70.10)

Demande n° Brevet n°	Date de publication (jour/mois/année)	Date de dépôt (jour/mois/année)	Date de priorité (valablement revendiquée) (jour/mois/année)
WO 97/28261	7.8.1997	31.1.1997	31.1.1996

L'attention du demandeur est attirée sur le fait que le document mentionné ci-dessus, au bénéfice d'une date de priorité antérieure à la présente demande, pourra être utilisé dans la phase régionale européenne pour autant que son contenu soit pertinent.

Cet examen est fondé sur l'hypothèse que toutes les revendications jouissent d'un droit de priorité datant du dépôt du document de priorité. S'il s'avérait qu'il n'en était pas ainsi, le document suivant, Nature, vol. 388, 1997, p. 882-887, cité dans le rapport de recherche internationale pourrait devenir pertinent.

Section VII

1. Contrairement à ce qu'exige la règle 5.1 a) ii) PCT, la description n'indique pas l'état de la technique antérieure pertinent exposé dans les documents D1-D3 et ne cite pas ces documents.
2. La référence à la figure 12b à la page 52 n'est pas correcte (cf. pages 43 et 52).
3. Le numéro d'application cité à la page 1, ligne 22 contrevient aux Directives C-II, 4.17 PCT.

Section VIII

1. Les revendications de la présente demande ne sont pas claires (Art. 6 PCT). Les protéines fluorescentes objet de la présente invention sont les protéines de cnidaires présentant les paramètres définis e.g. dans la revendication 1 (plus de $14000\text{ M}^{-1}.\text{cm}^{-1}$ et plus de 0,38 de rendement quantique de fluorescence). La protéine BFP est, d'une part exclue de cette définition (cf. description de la demande, page 11, ligne 11) et d'autre part, fait partie de l'invention (cf. description de la demande, page 4, ligne 21 et page 33, ligne 8 et voir Directives C-III, 4.3 PCT).
2. Les directives PCT, C-III, 2.1 stipulent que les revendications doivent exposer les caractéristiques techniques de l'invention. Les références à des publications scientifiques ne sont pas des caractéristiques techniques (e.g. revendication 26). De plus, dans les revendications 8, 14 et 25 l'emploi du terme notamment rend les références optionnelles et celles-ci n'ont pas de portée limitative sur lesdites revendications (Art. 6, Règle 6 PCT).
3. Les revendications 1-3, 5, 6, 22-24 ne sont pas claires pour les raisons suivantes (Art. 6 PCT). La partie introductory desdites revendications fait référence à des protéines fluorescentes de cnidaires définies par leur coefficient d'extinction moléculaire supérieur à $14000\text{ M}^{-1}.\text{cm}^{-1}$ et le rendement quantique de fluorescence supérieur à 0,38. L'emploi du terme notamment n'ayant pas de

fonction limitative sur lesdites protéines fluorescentes, il s'ensuit que lesdites revendications **ne sont pas limitées** à la GFP, contrairement à ce qu'il semble être suggéré dans la partie consacrée à la détection des dites revendications. Ainsi, les revendications comportent une contradiction qui ne permet pas de déterminer leur étendue.

4. La caractéristique de la revendication 30, selon laquelle des cellules ou fragments de cellules comprenant le gène codant pour une substance fluorescente fusionnée avec le gène d'un ligand sont revendiquées, n'est pas mentionnée dans la description. La revendication 30 ne se fonde donc pas sur la description, comme l'exige l'article 6 PCT.
5. L'emploi de parenthèses et crochets dans les revendications doit être limité aux cas où cela est absolument nécessaire (cf. Directives, C-III, 4.11 PCT). Il n'est pas clair si ces parenthèses et crochets sont limitatifs pour les revendications en contenant (Article 6 PCT).
6. La revendication 7 ne satisfait pas aux conditions énoncées dans l'article 6 PCT. Deux types différents de protéine fluorescente sont utilisés à choix dans la revendication 7, soit EGFP ou ECFP. Cependant, la partie introductory de cette revendication ne fait pas mention de la protéine fluorescente ECFP.
7. Les revendications 15 et 16 ne sont pas claires (article 6 PCT). Ces deux revendications font mention d'une substance marquée. Ces revendications ne décrivent cependant pas la substance marquée mais le marqueur seul (Bodipy et coumarine).
8. La revendication 32 n'est pas claire (article 6 PCT). La trousse revendiquée contient apparemment une protéine cible couplée à une protéine fluorescente quelconque (cf. page 79, ligne 7). Cependant, la protéine fluorescente décrite auparavant dans la revendication est choisie parmi des protéines fluorescentes de cnidaires ayant les propriétés optiques indiquées.
9. Les termes notamment, etc..., tel/telle que, et éventuellement utilisés dans les revendications 1, 2, 4-8, 10, 11, 13, 14, 17-23, 25, 27, 28, 31 et 32 sont vagues et

RAPPORT D'EXAMEN

Demande internationale n° PCT/FR98/01136

PRELIMINAIRE INTERNATIONAL - FEUILLE SEPARÉE

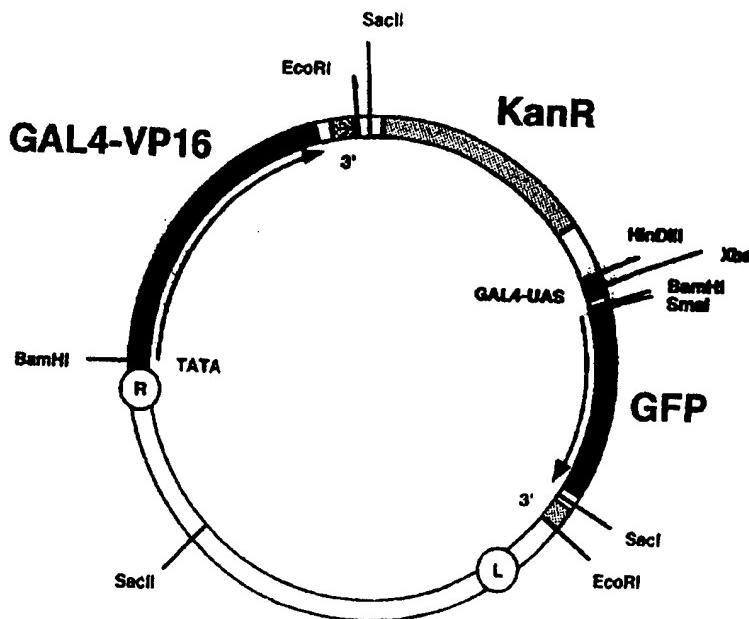
équivoques, et laissent un doute quant à la signification des caractéristiques techniques auxquelles ils se réfèrent. La même remarque s'applique au terme environ employé pour définir le coefficient d'extinction moléculaire supérieur à $14000\text{ M}^{-1}.\text{cm}^{-1}$ et le rendement quantique de fluorescence supérieur à 0.38 des protéines autofluorescentes de cnidaires dans les revendications 1, 2, 4, 6, 10, 11, 21-24, 27, 29-32. L'objet desdites revendications n'est donc pas clairement défini (article 6 PCT et directives C-III, 4.5a et 4.6 PCT).



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/82, A01H 5/00	A1	(11) International Publication Number: WO 97/30164 (43) International Publication Date: 21 August 1997 (21.08.97)
(21) International Application Number: PCT/GB97/00406 (22) International Filing Date: 14 February 1997 (14.02.97)		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).
(30) Priority Data: 9603069.7 14 February 1996 (14.02.96) GB		
(71) Applicant (for all designated States except US): MEDICAL RESEARCH COUNCIL [GB/GB]; 20 Park Crescent, London W1N 4AL (GB).		
(72) Inventors; and (73) Inventors/Applicants (for US only): HASELOFF, James, Phillip [AU/GB]; Glebe Cottage, Church Lane, Comberton, Cambridgeshire CB3 7ED (GB). HODGE, Sarah [GB/GB]; 59 Panton Street, Cambridge CB2 1HL (GB).		
(74) Agent: KEITH W. NASH & CO.; 90-92 Regent Street, Cambridge CB2 1DP (GB).		

(54) Title: IMPROVEMENTS IN OR RELATING TO GENE EXPRESSION



(57) Abstract

Disclosed is a nucleic acid sequence, expressible in a plant cell, encoding at least an effective portion of a GAL4 DNA-binding domain, the sequence having a % A/T base content substantially reduced relative to the wild-type yeast sequence.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LJ	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

Title: Improvements in or Relating to Gene Expression

Field of the Invention

This invention relates to a nucleic acid sequence optimised for expression in a plant cell, and to vectors and plant cells comprising the nucleic acid sequence.

Background of the Invention

In yeast there is a gene called GAL4, the product of which acts as a transcriptional activator (Johnston 1987 Microbiol. Rev. 51, 458-476). The GAL4 gene is quite large (3Kb) and the encoded polypeptide comprises an N-terminal DNA-binding domain, a C terminal domain having a transcriptional activator function, and an intervening glucose-responsive element "GRE" (GAL4 being repressed in the presence of glucose).

The GAL4 protein has been quite well characterised. Amino acid residues 1-147 of the protein bind to DNA in a sequence-specific manner (Keegan *et al* 1986 Science 231, 699-704). The transcriptional activation function is associated with two short portions of the C terminal domain (Ma & Ptashne 1987 Cell 48, 847-853). The DNA sequence to which GAL4 binds has been identified as a 17mer, which must be present as a repeat for optimal GAL4 binding (Giniger *et al*, 1985 Cell 40, 767-774), and may be referred to as a GAL4-responsive "upstream activation sequence" (UAS): GAL4 binds to the UAS 5' of a gene by means of the DNA-binding domain, and the C-terminal domain causes up-regulation of transcription of the gene.

Recently, a two-element system has been developed for directing gene expression in *Drosophila melanogaster*. Brand and Perrimon (1993 Development 118: 401-415,) randomly inserted the gene encoding the GAL4 into the *Drosophila* genome using a P-

element based vector.

A large number of stable *Drosophila* lines were generated which each express GAL4 in a particular pattern, dependent on adjacent genomic DNA sequences. A chosen target gene could then be cloned under the control of GAL4 upstream activation sequences (UAS), separately transformed, and maintained silently in the absence of GAL4. Genetic crossing of this single line with any of the library of GAL4-containing lines allowed activation of the target gene in many different tissue and cell types, and the phenotypic consequences of mis-expression, including those lethal to the organism, could be conveniently studied. In addition, the library of GAL4-containing flies has become an increasingly characterised and shared resource, and which provides a common entry point for various "reverse" genetic techniques.

A similar system operable in plants would be highly desirable. However, expression of heterologous eukaryotic genes in plants has proved highly problematical in the past (see, for example, Green Fluorescent Protein) and efficient expression of GAL4 appears equally difficult, and it has been suggested that this is due to inefficient translation of GAL4 mRNA in plants (Reichel *et al*, 1995 Plant Cell Reports 14, 773-776). Ma *et al*, (1988 Nature 334, 631-633) were able to obtain transient expression of modified, functional GAL4 derivatives in tobacco-leaf protoplasts but could not demonstrate the presence of functional, full-length wild-type GAL4. Similarly, transient expression of GAL4 derivatives has been demonstrated in maize protoplasts (McCarty *et al*, 1991 Cell 66, 895-905) and when introduced by biolistic methods, in maize aleurone tissues or embryogenic calli (Goff *et al*, 1991 Genes & Dev. 5, 298-309; Goff *et al*, 1992 Genes & Dev. 6, 864-875). Hitherto however, there have been no reports of stable, efficient expression of functional GAL4 or derivatives thereof in a plant cell.

It is an aim of the present invention to provide a novel expression system operable in plants, utilising a modified portion of the GAL4 gene.

Summary of the Invention

In a first aspect the invention provides a nucleic acid sequence, expressible in a plant cell, encoding at least an effective portion of a GAL4 DNA-binding domain, the sequence having an A/T base content substantially reduced relative to the wild-type yeast sequence.

The A/T content of the wild-type yeast sequence encoding the DNA-binding domain of GAL4 is about 59%. The % A/T base content of the sequence of the invention encoding the effective portion of the GAL4 DNA-binding domain will be understood to be substantially reduced when it is less than 50%. Preferably the A/T content is less than 45%, and more preferably less than 40%. The sequence of the invention may be made, for example, by site-directed mutagenesis, or be made *de novo* by chemical synthesis.

An "effective portion" of the DNA-binding domain is a portion sufficient to retain most (i.e. over 50%) of the DNA-binding activity of the full length DNA-binding domain. Typically the "effective portion" will comprise at least two thirds of the full length sequence of the DNA-binding domain. Conveniently the nucleic acid sequence will encode substantially all of the GAL4 DNA-binding domain (i.e. amino acid residues 1-147 of the yeast polypeptide), although a substantially smaller portion (about 75 amino acid residue) is quite adequate to retain most of the DNA-binding activity (Kraulis *et al*, 1992 *Nature* 356, 448-450; and Marmorstein *et al*, 1992 *Nature* 356, 408-414). In one particular embodiment, the sequence will comprise the nucleotide sequence shown in the 5' portion (nucleotides 1 to about 460) of Figure 1, which sequence has a substantially reduced A/T content (38%), relative to the wild-type yeast sequence, yet encodes an identical amino acid sequence.

The GAL4 DNA-binding domain has no transcriptional activation function in its own right. Thus, in preferred embodiments, the sequence encoding the effective portion of the GAL4 DNA-binding domain will be operably linked to one or more other nucleic acid sequences, which sequences may be structural (i.e. encode functional polypeptides) and/or regulatory. Typically the sequence encoding the DNA-binding domain will be operably linked (e.g. fused in-frame) to a sequence encoding a peptide or polypeptide with a regulatory function, preferably a transcriptional activator. The transcriptional activator may be the activation domain of GAL4 protein, the sequence encoding which should

preferably be optimised for expression in plants (e.g. by reducing the A/T content thereof). Alternatively, the transcriptional activator could be any one of a number of such proteins known to be active in plants, which will be well-known to those skilled in the art, such that the sequence of the invention encodes a chimeric polypeptide, comprising at least an effective portion of the GAL4 DNA-binding domain and a transcriptional activation domain. A particularly suitable transcriptional activator domain is that obtainable from herpes simplex virus (HSV) VP-16, (see Greaves and O'Hare 1989 *J. Virol.* 63, 1641-1650; VP, being also known as VMW65). Other transcriptional activation domains include certain peptides encoded by *E. coli* genomic DNA fragments (Ma & Ptashne 1987 *Cell* 51, 113-119) or synthetic peptides designed to form amphiphilic α -helix (Giniger & Ptashne 1987 *Nature* 330, 670-672). A common feature appears to be the requirement for excess charge, either positive (Gill & Ptashne 1987 *Cell* 51, 121-126) or, more especially for plant activation domains, excess negative charge (Estruch *et al.*, 1994 *Nucl. Acids Res.* 22, 3983-3989). With this in mind, the person skilled in the art could readily synthesise, or find naturally-occurring sequences, which encode peptides or polypeptides with transcriptional activation activity in plants. Such activator sequences may, in any event, be modified for optimal activity in a plant cell.

In a further aspect the invention provides a nucleic acid construct, comprising the nucleic acid sequence defined above. In one particular embodiment, the nucleic acid construct could be used as an "enhancer-trap" to "fish" for plant enhancer sequences (Cf. Sundaresan *et al.*, 1995 *Genes & Dev.* 9, 1797-1810). In such an embodiment, the construct will preferably include right and left Ti-DNA, to allow for random, stable insertion into the genome of a plant cell host. The construct will preferably comprise a naive plant promoter sequence, by which term is meant a promoter (operable in a plant cell) which requires the presence of a suitable enhancer sequence to cause substantial levels of transcription. Such a "naive" promoter may be thought of as an "enhancer-dependent" promoter, and essentially corresponds to the TATA box region of known plant promoters. "Plant" promoters includes reference to viral and bacterial promoter sequences which are active in a plant cell (e.g. nucleotides -48 to +1 of the CaMV 35S promoter). The naive promoter is conveniently substantially adjacent to one Ti border, the promoter being in competent relationship with the sequence encoding the GAL4 DNA-binding

domain fused with a transcription activation domain operable in plants such that, should the promoter become inserted in the plant host cell genome in functional relationship with an enhancer sequence, the promoter will direct the expression of the GAL4 DNA-binding domain and the transcriptional activation domain. The construct may additionally comprise a plant selectable marker (e.g. Kanamycin resistance) for convenience.

Additionally, the construct may comprise a reporter gene active in plants (such as Green Fluorescent Protein, or β -glucuronidase "GUS") in operable linkage to a GAL4-responsive upstream activation sequence (UAS), such that the reporter gene will be expressed in response to synthesis of the GAL4 DNA-binding domain/transcriptional activator fusion protein. Preferably the reporter gene is modified GFP as disclosed in WO 96/27675, or GUS (conveniently with a nuclear localisation signal). Preferably the UAS comprises one or more repeats of the 17mer to which GAL4 binds (Giniger *et al.*, 1985 Cell 40, 767-774).

As an alternative method of introducing the GAL4 binding domain/transcriptional activator sequence into a plant host cell genome, the construct may comprise Ds ("dissociation") elements, which may then be moved about the genome by the action of a transiently-expressed Ac (Activator) enzyme (e.g. Cocherel *et al.*, 1996 Plant Mol. Biol. 30, 539-551).

In a further aspect, the invention provides a plant or part thereof (e.g. a plant host cell or cell line) comprising the construct defined above. Typically the construct will have become integrated into a plant cell genome and be stably maintained therein. In particular the invention provides a plurality of plants or parts thereof (such as isolated plant cells, or cell lines, or tissue cultures) comprising a library, each plant or part thereof comprising a stably maintained nucleic acid sequence encoding an effective portion of the GAL4 DNA-binding domain as defined above. Conveniently the nucleic acid sequence will be incorporated into the genome of plant cells present in the library.

The library of plants or parts thereof thus provides a very useful resource. The library may be e.g. *Arabidopsis thaliana*, or of any other plant which is routinely used in

research. Each plant or part thereof in the library may have a particular pattern of expression of the integrated reporter gene (see, for example, Sundaresan *et al.*, 1995 *Genes & Dev.* 9, 1797-1810; Klimyuk *et al.*, 1995 *Molec. Gen. Genet.* 249, 357-365). Thus, introduction of a further gene, having a GAL4-responsive UAS into the cell line, will result in expression of the introduced gene in the same temporal/spatial pattern as the reporter gene, such that researchers can express a particular gene of interest in selected tissues and/or at selected times in a predictable manner.

The invention thus provides a method of expressing a gene of interest in a known (predictable) pattern in a plant or part thereof, the method comprising: introducing the gene of interest into the plant or part thereof, said gene of interest having a GAL4-responsive upstream activation sequence (UAS), characterised in that said plant or part thereof comprises a reporter gene expressed in a known pattern under the influence of a transcriptional activator comprising an effective portion of a GAL4 DNA-binding domain encoded by a sequence in accordance with the invention defined above, such that binding of the transcriptional activator to the UAS causes transcriptional activation of the gene of interest.

The temporal/spatial pattern of expression of the introduced gene or genes of interest will then mirror that already known for the reporter gene. Typically the plant or part thereof will be *Arabidopsis thaliana*, but may be any other plant which is routinely studied (e.g. maize, rice, potato etc).

Typically the gene or genes of interest will be introduced by crossing GAL4-expressing plant cell lines with a plant cell line comprising the gene(s) of interest linked to a GAL4-responsive UAS.

The nucleic acid sequence of the invention has other useful applications, in addition to use as an "enhancer trap", or to direct the expression of a gene or genes of interest in a predictable manner. The modified GAL4 DNA-binding domain/transcriptional activator may also be used to co-ordinate the expression of a plurality of genes of interest. In many circumstances it is desirable to express simultaneously a number of genes in a plant either

for research purposes, or for agricultural/industrial motives (e.g. to study metabolic pathways, or to manipulate the synthesis of desirable products such as plant dyes or lipids).

Thus, the invention provides a method for co-ordinating the expression of a plurality of genes of interest in a plant or part thereof, each of the genes of interest being functionally associated with a GAL4-responsive UAS, characterised in that the plant or part thereof comprises a sequence in accordance with the first aspect of the invention defined above and is capable of expressing a transcriptional activator comprising an effective portion of a GAL4 DNA-binding domain, such that binding of the transcriptional activator to the UAS causes simultaneous transcriptional activation of all of the genes of interest.

Conveniently the plurality of genes of interest may all be associated, in a polycistronic arrangement, with a single UAS, which facilitates their introduction into the plant or part thereof (the genes of interest typically being genes which are not naturally found in the plant or part thereof e.g. mammalian or bacterial genes, or genes from a different plant or possibly natural plant genes which have been the subject of various modifications). Alternatively one or more genes may be operably linked to a respective UAS. The transcriptional activator may be encoded by a sequence already present in the plant or part thereof, or may be introduced simultaneously with (or subsequent to) introduction of the plurality of genes of interest.

The invention will now be further described by way of illustrative examples and with reference to the accompanying drawings, of which:

Figure 1 shows a nucleotide sequence in accordance with the invention encoding a modified GAL4 DNA-binding domain, fused in frame to a sequence encoding a modified transcriptional activation domain from HSV VP16, the vertical line denotes the end of the GAL4 DNA-binding domain;

Figure 2 is a schematic representation of a nucleic acid construct comprising a sequence in accordance with the invention;

Figure 3 shows the nucleotide sequence of part of the nucleic acid construct represented schematically in Figure 2; and

Figures 4a-4d show micrographs of roots of plants comprising a sequence in accordance with the invention.

Examples

The inventors have used an approach similar to that described by Brand & Perrimon (1993, cited above), but adapted it for use in *Arabidopsis*, using *Agrobacterium tumefaciens* - mediated transformation to produce "enhancer trap" plant cell lines with a transcription activator comprising the GAL4 DNA-binding domain, fused to the activation domain of HSV VP16 (mGAL4-VP16).

Example 1 - Construction of mGAL4-VP16 gene

Eight oligonucleotides, which correspond to the modified sequence of the GAL4 DNA binding region, were synthesised.

GAL-5' (28-mer)

GGC AAC AAT GAA GCT ACT GTC TTC TAT C (Seq. ID No. 3)

VP16-3' (31-mer)

GGC AGA TCT ACC CAC CGT ACT CGT CAA TTC C (Seq. ID No. 4)

MGAL1 (109-mer)

GGC AAG CTT GGA TCC AAC AAT GAA GCT CCT GTC CTC CAT CGA GCA
GGC CTG CGA CAT CTG CCG CCT CAA GAA GCT CAA GTG CTC CAA GGA
GAA GCC GAA GTG CGC CAA G (Seq. ID No. 5)

MGAL2 (108-mer)

TCC TCT CGA GGG AAG ATC AGG AGG AAG AGC TGC TCC AGG CGC TCC

AGG CGG GAC TCC ACT TCG GTG AGG TGG GCG CGG GTC AGC GGG GAG
CGC TTG GTT TTG GGA GAG (Seq. ID No. 6)

MGAL3 (81-mer)

TTC CCT CGA GAG GAC CTC GAC ATG ATC CTG AAA ATG GAC TCC CTC
CAG GAC ATC AAA GCC CTG CTC ACC GGC CTC TTC GTC (Seq. ID No. 7)

MGAL4 (89-mer)

GGG TGA GGG GCA TGT CGG TCT CCA CGG AGG CCA GGC GGT CGG TGA
CGG CGT CTT TGT TCA CGT TGT CCT GGA CGA AGA GGC CGG TGA GC (Seq.
ID No. 8)

MGAL5 (80-mer)

GGA GAC CGA CAT GCC CCT CAC CCT GCG CCA GCA CCG CAT CAG CGC
GAC CTC CTC GGA GGA GAG CAG CAA CAA GGG CC (Seq. ID No. 9)

MGAL6 (83-mer)

AGT GGA GCT CGT CCC CCA GGC TGA CGT CGG TCG GGG GGG CCG TCG
AGA CGG TCA ACT GGC GCT GGC CCT TGT TGC TGC TCT CC (Seq. ID No. 10)

The full-length oligonucleotides were purified by electrophoresis in a 5% polyacrylamide gel containing 7M urea and 90mM Tris-borate 1mM EDTA pH8.3 (TBE) buffer. The fractionated oligonucleotides were detected by brief staining with filtered 0.05% toluidine blue dye, excised and eluted overnight at 37°C in 0.5M ammonium acetate, 0.1% SDS and 0.1mM EDTA. The eluted DNAs were precipitated and washed with ethanol, and phosphorylated using T4 polynucleotide kinase. DNAs (0.5ug each) were separately resuspended in 50mM Tris-HCl pH 9.0, 10mM MgCl₂, 10mM DTT with 1mM ATP and 5 units T4 polynucleotide kinase, and incubated for 30 mins at 37°C.

The plasmid pCMVGal65 (Cousens *et al*, 1989 EMBO J. 8, 2337-2342) was used as a source of the GAL4-VP16 sequence with unmodified codon usage. Two restriction endonuclease fragments were isolated from the plasmid. A *SacI-SacI* fragment, which

contains the sequence encoding the DNA binding element of the *Saccharomyces cerevisiae* GAL4 protein, and a *SacI-KpnI* fragment, which encodes the activation domain of the herpes simplex virus VP16 protein, were purified by fractionation and elution from a low gelling temperature (LGT) 1% agarose gel.

The 5' portion of the *mGAL4-VP16* sequence was amplified by the polymerase chain reaction (PCR), using the GAL-5' and MGAL2 oligos as primers and the *SacI-SacI* GAL4 DNA fragment as a template. Primers and template were incubated under standard conditions with VENT DNA polymerase (New England Biolabs) and subjected to 30 cycles of 94°C for 30 secs, 50°C for 1 min, 72°C for 1 min. The product was cut with *XhoI* restriction endonuclease and purified after electrophoresis through a 1.5% LGT agarose gel.

The MGAL4 and MGAL5 oligonucleotides were annealed by heating to 65°C and slow cooling to room temperature. The MGAL3 and MGAL6 oligonucleotides were then added and also annealed to the mixture. The oligonucleotides were then incubated with Klenow fragment of DNA polymerase 1 and T4 DNA ligase in 50mM Tris-HCl pH 7.5, 10mM MgCl₂, 10mM DTT, 1mM dNTPs and 1mM ATP for 60 mins at 37°C. An excess of MGAL3 and MGAL6 oligonucleotides were added, and the mixture was subject to PCR amplification, as above. The product was cut with the restriction endonucleases *XhoI* and *SacI*, and purified after electrophoresis through a 1.5% LGT agarose gel.

The three gel-purified restriction fragments: (1) the GAL-5'/MGAL2 PCR-amplified product, corresponding to the 5' portion of the GAL4 DNA binding domain, (2) the MGAL3-6 annealed and amplified product, corresponding to the 3' portion of the GAL4 DNA binding domain, and (3) the *SacI-KpnI* fragment from pCMVGal65, corresponding to the VP16 activation domain; were mixed and ligated with T4 DNA ligase in 50mM Tris-HCl pH 7.5, 10mM MgCl₂, 10mM DTT and 1mM ATP at 20°C overnight. The ligated product was PCR amplified (as above) using MGAL1 and VP16-3' as primers. The product, containing the entire modified GAL4-VP16 sequence was cut with *BamHI* and ligated to PBI121 (Jefferson *et al.*, EMBO J. 6:3907, 1987) which had been cut with *SacI*, blunt ended by treatment with T4 DNA polymerase, and re-cut with *BamHI*. The

recombinant plasmid (pBIN 35S-*mGAL4-VP16*) contained the *mGAL4-VP16* gene downstream of the constitutive 35S plant promoter.

In Figure 1, the lower DNA sequence is that of the fusion gene (Seq. ID No. 1), encoding in the 5' portion the GAL4 DNA-binding domain, and encoding in the 3' portion the transcriptional activation domain from HSV VP16. The wild-type sequence is shown above for comparison. It will be seen that the wild type GAL4 DNA-binding domain coding sequence is A/T rich (A/T% content is shown on the right for each line of sequence), and it is believed that this causes inefficient expression in plants. In particular, it is believed that the A/T rich DNA comprises one or more regions which, when transcribed, are recognized as mRNA splice sites in plant cells. The altered nucleotides are shown outlined in Figure 1. The encoded polypeptide sequence is shown below (Seq. ID No. 2). The modifications to the nucleotide sequence were carefully selected to ensure that there was no resulting change in the encoded amino acid sequence. The numbers above the sequences represent the number of the nucleotide or amino acid at that position. Certain restriction endonuclease sites are also shown.

The 35S promoter and *mGAL4-VP16* gene were excised by digestion with the restriction endonucleases *Hin*DIII and *Eco*RI, purified by 1% LGT agarose gel electrophoresis, and subcloned into a vector containing a GAL4-responsive *mgfp5-ER* marker gene (see Example 2 below). The resulting construct is shown in Figure 2, in which R and L represent right and left Ti plasmid border sequence. The GAL4-VP16 fusion gene is downstream of the TATA-box region (nucleotides -48 to +1) of the CaMV 35S gene, were inserted adjacent to the right Ti border. The plasmid also contains the kanamycin resistance (Kan R) selectable marker, and a reporter gene (GUS or GFP) operably linked to a synthetic GAL4-UAS promoter sequence. Certain restriction endonuclease sites are also shown.

Figure 3 shows the nucleotide sequence (Seq. ID No. 15) of the plasmid across the junctions of the Ti right border, the CaMV35S TATA box, and the start of the GAL4 DNA-binding domain coding region.

The *mGAL4-VP 16* gene was directly assayed for activity in transformed *Arabidopsis* plants by *Agrobacterium*-mediated transformation (Valvekens *et al* Proc. Natl. Acad. Sci. USA 85:5536-5540, 1988).

Example 2

Construction of GAL-GFP enhancer trap vector.

Modification of the T-DNA right border.

The plasmid pBIN19 (Bevan, M. Nuc. Acids Res. 12:8711-8721, 1984) was digested with the restriction endonuclease *Sac*II, and a 2.6 Kb fragment containing the *Agrobacterium* T-DNA right border sequence was subcloned into *Sac*II cut pGEM 5Zf(Genbank accession No. X65308; commercially available from Promega). The bacteria containing the recombinant phagemid were superinfected with helper phage to produce single strand phagemid DNA.

A synthetic mutagenic oligonucleotide ("TR+": ATA TCC TGT CAA ACA CTG GAT CCG AGC TCC AAT TCA TAG TTT AAA CTG AAG GCG GG; Seq. ID No. 11) was kinased and annealed to the purified phagemid DNA, extended on the template using T7 DNA polymerase, and ligated using T4 DNA ligase. The DNA was transformed into bacteria, and recombinant plasmids (pTr+) were screened for the insertion of new sites for the restriction endonucleases *Bam*HI, *Sma*I and *Eco*RI immediately adjacent to the T-DNA right border.

Insertion of TATA box region.

The TATA box region of the 35S promoter was PCR amplified from pBI121 using oligonucleotides complementary to the -48 region of the promoter (oligonucleotide Δ 35SB g II; GGC AGA TCT TCG CAA GAC CCT TCC TCT ATA TAA GG; Seq. ID No. 12) and the downstream β -glucuronidase gene (oligonucleotide GUSSnaBI; CAC ACA AAC GGT GAT ACG TA; Seq. ID No. 13). The PCR product was cut with restriction endonucleases *Bg*II and *Bam*HI and ligated to *Bam*HI cut pTr+. The resulting recombinant plasmid (pTr+ Δ 35S) contained a minimal (naive) promoter positioned

adjacent to the T-DNA right border.

Insertion of 35S promoter driven *mGAL4-VP 16* gene.

Derivatives of the *GAL4* with unmodified codon usage were inserted into the pTr+Δ35S plasmid in unsuccessful trial experiments, and one of these constructions (the insertion of a truncated *GAL4* gene derived from pMA236; Ma *et al*, 1988 *Nature* 334, 631-633) had resulted in the insertion of a unique *HindIII* site immediately adjacent to the *BamHI* site in pTR+Δ35S. This plasmid, pTr+Δ*GAL4*, was digested with restriction endonucleases *HindIII* and *EcoRI*, and ligated with the *HinDIII-EcoRI* fragment from pBIN 35S-*mGAL4-VP 16*, to introduce the 35S-driven *mGAL4-VP 16* gene. The *Agrobacterium* sequences from this plasmid (pTr+35S-*mGAL4-VP 16*) were then recloned into a binary plant transformation vector which contained a *GAL4*-responsive marker gene (see below), for testing *in planta*.

Construction of a synthetic *GAL4*-responsive promoter.

A *HinDIII-XbaI* restriction endonuclease fragment containing five optimised binding sites for *GAL4* was excised from the plasmid pUAST (Brand & Perrimon, 1993 Development, 118, 401-415), and ligated into *HinDIII-XbaI* cut pBI101 (Jeferson *et al.*, 1987 EMBO J. 6, 3901-3907) upstream of the GUS gene (to make pBINΔUAS-GUS). A sequence corresponding to the -90 region of the 35S promoter was PCR amplified using oligonucleotides DELAS1 (GAA CTC TAG AAG CTA CTC CAC GTC CAT AAG GGA CAC ATC ACA ATC CCA CTA TCC TTC GC; Seq. ID No. 14) and GUSSnaBI (see above). The product was *XbaI-BamHI* cut and cloned into similarly cut pBIN ΔUAS-GUS. The resulting plasmid (pBIN UAS(-90-AS1)GUS) contains a synthetic *GAL4* promoter upstream of the β-glucuronidase (GUS) coding sequence. The GUS gene was then replaced with that of the green fluorescent protein (GFP).

GAL4-responsive *mgfp5-ER* gene.

We have extensively mutated the green fluorescent protein gene for optimised expression in plants. The optimised gene is called *mgfp5-ER* (Siemering *et al*, Current Biology 6:1653-1663, 1996; WO96/27675).

The *mgfp5-ER* gene has been cloned into the plant transformation vector pBI121, and we have excised a *Bam*HI-*Sac*I fragment, which contains the coding sequence, and have inserted this into *Bam*HI-*Sac*I cut pBIN UAS(-90-AS1)GUS to replace the GUS gene. This plasmid (pBIN UAS(-90-AS1)*mgfp5-ER*) was cut with *Sac*II, and ligated with the *Sac*II fragment containing the modified T-DNA right border and mGAL4-VP16 gene from pTr+35S-*mGAL4-VP16*, to give pBIN 35S-*GAL4-VP16+* UAS-*mgfp5-ER* (abbreviated to pBIN35S-GAL-GFP).

Testing of the pBIN35S-GAL-GFP vector.

The pBIN 35S-GAL-GFP plasmid was introduced into *Agrobacterium tumefaciens* strain LBA4044 (Jefferson *et al.*, EMBO J. 6:3901-3907, 1987) by electroporation, followed by selection on kanamycin-containing plates. The *Agrobacterium* strain was used to transform *Arabidopsis* using the technique described below. Plants were scored for the GAL4-dependent induction of GFP fluorescence after transformation.

Construction of GAL-GFP enhancer trap vector

The pTR+35S-*mGAL4-VP16* plasmid was treated with *Bam*HI (to excise the 35S promoter sequences) and religated using T4 DNA ligase. This plasmid (pTR+*mGAL4-VP16*) contains the modified *GAL4-VP16* positioned with a minimal promoter adjacent to the T-DNA right border. The T-DNA sequences were then excised by digestion with *Sac*II, and ligated with similarly cut pBIN UAS(-90-AS1)*mgfp5-ER*. This produced pBIN GAL-GFP, which contains the highly active *mGAL4-VP16* transcription activator, responsive to adjacent enhancer elements, and a GAL4-dependent GFP gene. GFP fluorescence is produced in response to GAL4-VP16 activation.

Enhancer trap screen.

The enhancer trap vector pBIN GAL-GFP was introduced into *Agrobacterium tumefaciens* (see details below) LBA4044, and has been used to generate over 7,500 transformed *Arabidopsis* seedlings. Transgenic seedlings have been screened directly for GAL4-mediated GFP expression, and seeds were collected from expressing plants. We have a library of over 250 *Arabidopsis* lines which show stable inheritable patterns of GAL4-

mediated GFP expression in the root.

Transactivation.

To demonstrate GAL4-mediated transactivation of a foreign gene we have generated transgenic *Arabidopsis* that contain the T-DNA sequences from pBIN UAS(-90-AS1)GUS. In the absence of GAL4-VP16, these plants do not express the GUS gene. However, when genetically crossed with lines expressing the *mGAL4-VP16* gene, the GUS gene is activated in a pattern of expression which accurately reflects that of the *mGAL4-VP16* gene from the donor parent (see Example 3 below).

Example 3

A stable transformed line (J2302) of *Arabidopsis thaliana*, which formed part of the library described above, expressed modified GFP (under the influence of the mGAL4-VP16 activator) in the cells of the extreme root tip (see Figures 4a and 4b).

Figure 4a is a low power micrograph showing GFP-mediated epifluorescence (400nm excitation) at the root tip. Figure 4b is a higher magnification confocal micrograph (488nm excitation) showing the area of GFP-mediated fluorescence in greater detail.

The line J2302 was crossed (using standard techniques) with another *Arabidopsis* line which comprised a stable, silently-maintained GUS reporter gene operably linked to a GAL4-responsive UAS. The product of the cross, as expected, proceeded to express GUS, under the influence of the GAL4-VP16 transcriptional activator, in the same pattern as the GFP reporter gene in J2302 (i.e. at the extreme root tip). Suitably stained samples, imaged by brightfield microscopy, are shown in Figures 4c and 4d, and demonstrate GUS reporter gene activity (darkest staining) at the root tip.

This illustrates that the modified GAL4 DNA-binding domain sequence can

- (i) successfully be used in the construction of enhancer trap vectors;
- (ii) express genes of interest (as exemplified by GUS) in a predictable pattern; and
- (iii) co-ordinate the simultaneous expression of a plurality of genes of interest (as

exemplified by GFP and GUS).

Example 4

Transformation of *Arabidopsis thaliana*.

The methods employed are based on those originally disclosed by Valvekens *et al.* 1988.
Proc. Natl. Acad. Sci. 85, 5536-5540).

Media

Unless otherwise stated all procedures were carried out aseptically using sterile solutions and equipment. All media are used in standard plastic disposable petri dishes or, for later stages, Magenta GA 7 pots (Magenta Corp., USA). Hormones are dissolved in dimethyl sulfoxide (DMSO) as x1000 stocks. Hormones and antibiotics are added after autoclaving and cooling of the media to 65°C.

1. Germination medium (GM) comprised: 1x Murashige and Skoog salt mixture; 1% sucrose; 100 mg/l inositol; 1.0 mg/l thiamine; 0.5 mg/l pyridoxine; 0.5 mg/l nicotinic acid; 0.5 g/l 2-(N-morpholino)ethanesulfonic acid (MES); (adjust to pH 5.7 with 1M KOH); and 0.8% Difco Bacto agar (for solid GM medium).

1a. GM K50

As GM, but supplemented with: 50 mg/l kanamycin (Sigma)

2. Callus-inducing medium (CIM) comprised: 1x Gamborg's B5 medium; 2% glucose; 0.5 g/l MES (pH 5.7, adjusted with 1N KOH); 0.5 mg/l 2,4-D (Sigma); 0.05 mg/l kinetin (Sigma); and 0.8 % agar (for solid CIM medium).

3. Shoot-inducing medium with Vancomycin (SIM V750 K50) comprised: Gamborg's B5 medium; 2% glucose; 0.5 g/l MES (pH 5.7); 0.8% agar; 5 mg/l N6-(2-isopentenyl)adenine (2ip); 0.15 mg/l indole-3-acetic acid (IAA); 750 mg/l vancomycin; and 50 mg/l kanamycin.

3a. SIM V500 K50

As SIM, but supplemented with: 500 mg/l vancomycin; and 50 mg/l kanamycin.

4. Root-inducing medium (RIM) comprised: Gamborg's B5 medium; 2% glucose: 0.5 g/l MES (pH 5.7); 0.8% agar; 12.5 mg/l indole butyric acid (IBA); and 50 mg/l cefotaxime.

Growth of plants

- (i) Place seeds into a 15 ml polypropylene centrifuge tube.
- (ii) Add ethanol for 2 min. Remove ethanol with pipette.
- (iii) Replace with 5% commercial bleach (~0.25% available chlorine) containing one drop of NP40 per 50 ml. Leave for 15 min, shaking regularly.
- (iv) Wash seeds in sterile, distilled water for at least three times.
- (v) After last wash, add the sterile seeds to 100ml of GM in a conical flask. Grow with shaking for 2-4 weeks in a culture room at 20-25°C.

Agrobacterium-mediated transformation.

1. Day One

- (i) Using a scalpel, cut-off all the green (upper) parts of the plantlets to leave only the root systems.
- (ii) Lay out the roots on plates containing solid CIM. Gently press down on each bunch of roots to ensure that they are in contact with the surface of the agar.
- (iii) Incubate for 3 d in growth room (22°C, continual light).

2. Day Four

- (i) After ensuring that there is no visible sign of contamination, collect the callus-induced roots in an empty petri dish.
- (ii) Cut the roots into 0.5 cm explants.
- (iii) Add 3-5 ml of *Agrobacterium* culture which has been grown overnight at 28°C in Luria broth, and washed by centrifugation. Swirl with blunt-nosed forceps. Leave to co-cultivate for 2min.
- (iv) Blot the roots dry on double-thick sterile filter-paper (Whatman No.1) in a petri dish.
- (v) Place the explants onto solidified CIM medium in a petri dish, and gently press them onto the surface to ensure contact with the medium.

(vi) Incubate plates in the growth room for 2 d to allow co-cultivation of Agrobacterium.

3. Day Six

(i) Transfer the explants to a petri dish in 20-25 ml of water.

Agitate with blunt-nosed forceps to wash off the agrobacteria. The bathing medium should become somewhat turbid. Transfer the root pieces to a sieve and repeat the washing. Then lift the sieve out of buffer and drain. Press down the explants down onto the mesh to remove as much buffer and agrobacteria as possible. Push the semi-dried explants together into a mass.

(Custom-made, autoclavable and re-usable, sieves are used to hold the explants during *Agrobacterium* infections and washings of root explants. These sieves are made from 100 ml plastic three-cornered beakers and 100 um nylon mesh. The top 3-4 cm is cut off the beaker and the mesh held in place across the bottom of this piece by pushing a ring 2 cm high cut from just above the base of the beaker into the lower portion of the top part. The two pieces are sealed together by pushing a hot metal rod through them in several places.)

(ii) Blot small bundles of root material on double-thick sterile filter-paper in a petri dish until dry and transfer to solidified SIM V750 K50 medium, taking care that root explants are in close contact with the medium.

(iii) Incubate in growth room.

Regeneration

(i) In growth room, tiny green kanamycin resistant calli appear on the yellowish root-explants after 3 weeks.

(ii) Every 2 weeks, transfer explants to fresh SIM V750 K50. Shoots (often initially vitreous) intermittently appear from the green calli over the next several weeks.

(iii) Transfer these shoots to RIM, where roots will develop over a period of 2 to 3 weeks

(iv) After the formation of roots, transfer the plantlets to GM in vented Magenta GA7 pots or directly to soil. Ensure that the humidity is kept low to ensure good seed set.

(v) Harvest seed pods when they turn yellowish-brown and are therefore mature.

Germination and screening of F1 progeny.

To test for transformation, F1 seedlings are germinated on GM K50.

- (i) Put seeds on GM K50 (surface sterilise the seeds if necessary).
- (ii) Put petri dishes in the dark at 4oC (refrigerator) for 3-5 d. to break seed dormancy. This is not necessary if seeds were stored for more than a month.
- (iii) Incubate petri dishes in growth room for 2 weeks. Sensitive seedlings form neither roots nor leaves, and have white cotyledons. Transformed seedlings are phenotypically normal.
- (iv) Transgenic callus and shoots were screened for GFP expression using a inverted fluorescence microscope (Leitz DM-IL) fitted with a filter set (Leitz-D excitation BP355-425, dichroic 455, emission LP460) suitable for the main 395nm excitation and 509nm emission peaks of GFP. The use of a 7mm threaded extension tube with a 4x objective (EF 4/0.12) gave a greater working distance, and has allowed the convenient direct observation of tissue within inverted sealed petri dishes. A 100 Watt long wavelength hand-held UV lamp (UV Products, B100AP) was also used for routine monitoring of transgenic shoots and plants. Transgenic *Arabidopsis* F1 seedlings were grown in sterile agar culture for 5 days, and were mounted in water under glass coverslips for microscopy. The specimens were examined using a BioRad MRC-600 laser-scanning confocal microscope equipped with a krypton-argon laser and filter sets suitable for the detection of fluorescein and texas red dyes (BioRad K1/K2), and a Nikon 60x PlanApo N.A. 1.2 water immersion objective.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: Medical Research Council
- (B) STREET: 20 Park Crescent
- (C) CITY: London
- (E) COUNTRY: United Kingdom
- (F) POSTAL CODE (ZIP): W1N 4AL
- (G) TELEPHONE: (0171) 636 5422
- (H) TELEFAX: (0171) 323 1331

(ii) TITLE OF INVENTION: Improvements in or Relating to Gene Expression

(iii) NUMBER OF SEQUENCES: 15

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 701 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 17..694

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

AAGCTTGGAT CCAACA ATG AAG CTC CTG TCC TCC ATC GAG CAG GCC TGC Met Lys Leu Leu Ser Ser Ile Glu Gln Ala Cys	49
1 5 10	
GAC ATC TGC CGC CTC AAG AAG CTC AAG TGC TCC AAG GAG AAG CCG AAG Asp Ile Cys Arg Leu Lys Leu Lys Cys Ser Lys Glu Lys Pro Lys	97
15 20 25	
TGC GCC AAG TGT CTG AAG AAC AAC TGG GAG TGT CGC TAC TCT CCC AAA Cys Ala Lys Cys Leu Lys Asn Asn Trp Glu Cys Arg Tyr Ser Pro Lys	145
30 35 40	
ACC AAG CGC TCC CCG CTG ACC CGC GCC CAC CTC ACC GAA GTG GAG TCC Thr Lys Arg Ser Pro Leu Thr Arg Ala His Leu Thr Glu Val Glu Ser	193
45 50 55	

CGC CTG GAG CGC CTG GAG CAG CTC TTC CTC CTG ATC TTC CCT CGA GAG Arg Leu Glu Arg Leu Glu Gln Leu Phe Leu Leu Ile Phe Pro Arg Glu 60 65 70 75	241
GAC CTC GAC ATG ATC CTG AAA ATG GAC TCC CTC CAG GAC ATC AAA GCC Asp Leu Asp Met Ile Leu Lys Met Asp Ser Leu Gln Asp Ile Lys Ala 80 85 90	289
CTG CTC ACC GGC CTC TTC GTC CAG GAC AAC GTG AAC AAA GAC GCC GTC Leu Leu Thr Gly Leu Phe Val Gln Asp Asn Val Asn Lys Asp Ala Val 95 100 105	337
ACC GAC CGC CTG GCC TCC GTG GAG ACC GAC ATG CCC CTC ACC CTG CGC Thr Asp Arg Leu Ala Ser Val Glu Thr Asp Met Pro Leu Thr Leu Arg 110 115 120	385
CAG CAC CGC ATC AGC GCG ACC TCC TCC TCG GAG GAG AGC AGC AAC AAG Gln His Arg Ile Ser Ala Thr Ser Ser Glu Glu Ser Ser Asn Lys 125 130 135	433
GGC CAG CGC CAG TTG ACC GTC TCG ACG GCC CCC CCG ACC GAC GTC AGC Gly Gln Arg Gln Leu Thr Val Ser Thr Ala Pro Pro Thr Asp Val Ser 140 145 150 155	481
CTG GGG GAC GAG CTC CAC TTA GAC GGC GAG GAC GTG GCG ATG GCG CAT Leu Gly Asp Glu Leu His Leu Asp Gly Glu Asp Val Ala Met Ala His 160 165 170	529
GCC GAC GCG CTA GAC GAT TTC GAT CTG GAC ATG TTG GGG GAC GGG GAT Ala Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Asp Gly Asp 175 180 185	577
TCC CCG GGG CCG GGA TTT ACC CCC CAC GAC TCC GCC CCC TAC GGC GCT Ser Pro Gly Pro Gly Phe Thr Pro His Asp Ser Ala Pro Tyr Gly Ala 190 195 200	625
CTG GAT ATG GCC GAC TTC GAG TTT GAG CAG ATG TTT ACC GAT GCC CTT Leu Asp Met Ala Asp Phe Glu Phe Glu Gln Met Phe Thr Asp Ala Leu 205 210 215	673
GGA ATT GAC GAG TAC GGT GGG TAGATCT Gly Ile Asp Glu Tyr Gly Gly 220 225	701

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 226 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Lys Leu Leu Ser Ser Ile Glu Gln Ala Cys Asp Ile Cys Arg Leu
 1 5 10 15

Lys Lys Leu Lys Cys Ser Lys Glu Lys Pro Lys Cys Ala Lys Cys Leu
 20 25 30

Lys Asn Asn Trp Glu Cys Arg Tyr Ser Pro Lys Thr Lys Arg Ser Pro
 35 40 45

Leu Thr Arg Ala His Leu Thr Glu Val Glu Ser Arg Leu Glu Arg Leu
 50 55 60

Glu Gln Leu Phe Leu Leu Ile Phe Pro Arg Glu Asp Leu Asp Met Ile
 65 70 75 80

Leu Lys Met Asp Ser Leu Gln Asp Ile Lys Ala Leu Leu Thr Gly Leu
 85 90 95

Phe Val Gln Asp Asn Val Asn Lys Asp Ala Val Thr Asp Arg Leu Ala
 100 105 110

Ser Val Glu Thr Asp Met Pro Leu Thr Leu Arg Gln His Arg Ile Ser
 115 120 125

Ala Thr Ser Ser Ser Glu Glu Ser Ser Asn Lys Gly Gln Arg Gln Leu
 130 135 140

Thr Val Ser Thr Ala Pro Pro Thr Asp Val Ser Leu Gly Asp Glu Leu
 145 150 155 160

His Leu Asp Gly Glu Asp Val Ala Met Ala His Ala Asp Ala Leu Asp
 165 170 175

Asp Phe Asp Leu Asp Met Leu Gly Asp Gly Asp Ser Pro Gly Pro Gly
 180 185 190

Phe Thr Pro His Asp Ser Ala Pro Tyr Gly Ala Leu Asp Met Ala Asp
 195 200 205

Phe Glu Phe Glu Gln Met Phe Thr Asp Ala Leu Gly Ile Asp Glu Tyr
 210 215 220

Gly Gly
 225

(2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 28 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GGCAACAAATG AAGCTACTGT CTTCTATC

28

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GGCAGATCTA CCCACCGTAC TCGTCAATT C

31

(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 109 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GGCAAGCTTG GATCCAACAA TGAAGCTCCT GTCCTCCATC GAGCAGGCCT GCGACATCTG

60

CCGCCTCAAG AAGCTCAAGT GCTCCAAGGA GAAGCCGAAG TGCGCCAAG

109

(2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 108 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

T CCTCTCGAG GGAAGATCG GAGGAAGAGC TGCTCCAGGC GCTCCAGGCG GGACTCCACT

60

TCGGTGAGGT GGGCGCGGGT CAGCGGGGAG CGCTTGGTTT TGGGAGAG

108

(2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 81 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

TTCCCTCGAG AGGACCTCGA CATGATCCTG AAAATGGACT CCCTCCAGGA CATCAAAGCC	60
CTGCTCACCG GCCTCTTCGT C	81

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 89 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GGGTGAGGGGG CATGTCGGTC TCCACGGAGG CCAGGCGGTC GGTGACGGCG TCTTTGTTCA	60
CGTTGTCCTG GACGAAGAGG CGCGTGAGC	89

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 80 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GGAGACCGAC ATGCCCTCA CCCTGCGCCA GCACCGCATC AGCGCGACCT CCTCCTCGGA	60
GGAGAGCAGC AACAAAGGGCC	80

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 83 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

AGTGGAGCTC GTCCCCCAGG CTGACGTCGG TCGGGGGGGC CGTCGAGACG GTCAACTGGC	60
GCTGGCCCTT GTTGCTGCTC TCC	83

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 56 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

ATATCCTGTC AAACACTGGA TCCGAGCTCC AATTCTAGT TTAAACTGAA GGCGGG

56

(2) INFORMATION FOR SEQ ID NO: 12:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 35 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GGCAGATCTT CGCAAGACCC TTCCTCTATA TAAGG

35

(2) INFORMATION FOR SEQ ID NO: 13:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

CACACAAACG GTGATACGTA

20

(2) INFORMATION FOR SEQ ID NO: 14:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 59 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

GAACTCTAGA AGCTACTCCA CGTCCATAAG GGACACATCA CAATCCACT ATCCTTCGC

59

(2) INFORMATION FOR SEQ ID NO: 15:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 91 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

TAGGTTTACC CGCCAATATA TCCTGTAAA CACTGGATCT TCGCAAGACC CTTCCCTCTAT 60
ATAAGGAAGT TCATTCATT TGGAGAGGAC A 91

CLAIMS

1. A nucleic acid sequence, expressible in a plant cell, encoding at least an effective portion of a GAL4 DNA-binding domain, the sequence having a % A/T base content substantially reduced relative to the wild-type yeast sequence.
2. A sequence according to claim 1, further comprising a structural and/or regulatory sequence in operable linkage to the sequence encoding the effective portion of the GAL4 DNA-binding domain.
3. A sequence according to claim 1 or 2, wherein the nucleic acid encodes a peptide or polypeptide, having transcriptional activation activity, in operable linkage with the sequence encoding the effective portion of the GAL4 DNA-binding domain.
4. A sequence according to any one of claims 1, 2 or 3, wherein the sequence encodes a peptide or polypeptide, having transcriptional activation activity in a plant cell, which is a modified portion of GAL4, VP16, GCN4 or a designed sequence which does not naturally occur.
5. A nucleic acid construct comprising a sequence in accordance with any one of claims 1-4.
6. A construct according to claim 5, capable of inserting a sequence according to any one of claims 1-4 into the genome of a plant host cell.
7. A construct according to claim 5 or 6, further comprising a reporter gene in operable linkage with a GAL4-responsive upstream activation sequence (UAS).
8. A construct according to any one of claims 5, 6 or 7, comprising a sequence encoding GUS or green fluorescent protein (GFP) suitable for expression in a plant cell.

9. A construct according to any one of claims 5-8, comprising T-DNA borders or Ds elements.
10. A construct according to any one of claims 5-9, wherein expression of the sequence according to claims 1-4 is under the control of an enhancer-dependent plant promoter.
11. A construct according to any one of claims 5-10, further comprising a selectable marker gene.
12. A plant or part thereof comprising a nucleic acid sequence according to any one of claims 1-4, or transformed with a construct according to any one of claims 5-11.
13. A plant or part thereof according to claim 12, wherein the sequence is stably maintained in the plant genome.
14. A library comprising a plurality of plants or parts thereof, each plant or part thereof comprising a stably integrated nucleic acid sequence according to claim 3 or claim 4, which sequence causes the expression of a reporter gene in a known temporal and/or spatial pattern.
15. A library according to claim 14, made by the introduction of a nucleic acid construct according to any one of claims 6-11 into a plurality of plants or parts thereof.
16. A library according to claim 14 or 15, comprising a plurality of *Arabidopsis thaliana* plants.
17. A method of expressing a gene of interest in a known pattern in a plant or part thereof, the method comprising: introducing the gene of interest into the plant or part thereof, said gene of interest having a GAL4- responsive upstream activation sequence (UAS), characterised in that said plant or part thereof comprises a reporter gene expressed in a known pattern under the influence of a transcriptional activator comprising an effective portion of a GAL4 DNA-binding domain encoded by a sequence according to any

one of claims 1-4, such that binding of the transcriptional activator to the UAS causes transcriptional activation of the gene of interest.

18. A method according to claim 17, wherein the gene of interest is introduced into one or more plants or parts thereof from a library according to any one of claims 14, 15 or 16.

19. A method for co-ordinating the expression of a plurality of genes of interest in a plant or part thereof, each of the genes of interest being functionally associated with a GAL4-responsive UAS, characterised in that the plant or part thereof comprises a sequence according to any one of claims 1-4 and is capable of expressing a transcriptional activator comprising an effective portion of a GAL4 DNA-binding domain, such that binding of the transcriptional activator to the UAS causes simultaneous transcriptional activation of all of the genes of interest.

20. A method according to claim 19, wherein at least some of the genes of interest are arranged polycistronically and associated with a single GAL4-responsive UAS.

21. A method according to claim 19, wherein genes of interest are functionally associated with respective GAL4-responsive UASs.

A/T content											
<i>Hind III BamH I</i>											32/11
atg aag cta ctg tct atc gaa caa gca tgc gat att tgc ggatcc aaca atg aag ctc ctg tcc atc gag ca <u>g</u> gcc tgc gac at <u>c</u> tgc											60% 40%
M K L S I E Q A C D I C											
62/21	92/31										
cga ctt aaa aag ctc aag tgc tcc aaa gaa aaa ccg aag tgc gcc aag tgt ctg aag aac cg <u>c</u> ct <u>c</u> aa <u>g</u> aag ctc aag tgc tcc aa <u>g</u> gag ccg aag tgc gcc aag tgt ctg aag aac											55% 45%
R L K C S K E P K C A K C L K N											
122/41	152/51										
aac tgg gag tgt cgc tac tct ccc aaa acc aag cg <u>c</u> tcc ccg ctg acc cg <u>c</u> gcc cac ct <u>c</u> aac tgg gag tgt cgc tac tct ccc aaa acc aag cg <u>c</u> tcc ccg ctg acc cg <u>c</u> gcc cac ct <u>c</u>											47% 35%
N W E C R Y S P K T K R S P L T R A H L											
182/61	212/71										
aca gaa gtg gaa tca agg cta gaa aga ctg gaa cag cta ttu cta ctg att ttu cct cga acc gaa gtg gag tcc cg <u>c</u> ct <u>g</u> gag gg <u>c</u> ctg gag cag ct <u>c</u> ttu ct <u>c</u> ctg at <u>c</u> tt <u>c</u> cct cga											60% 37%
T E V S R L E R L E Q F L L I F P R											
242/81	272/91										
gaa gac ctt gac att ttg aaa atg gat ttu cag gat ata aaa gca ttg tta aca gag gac ct <u>c</u> gac at <u>c</u> ctg aaa atg gac tcc ct <u>c</u> gag gac at <u>c</u> aaa gcc ct <u>c</u> acc											70% 42%
E D L M I L K M D S L Q D I K A L L T											
302/101	332/111										
gg <u>a</u> ttu ttt gta caa gat aat gtg aat aaa gat gcc gtc aca gat aga ttg gct tca gtg gg <u>c</u> ct <u>c</u> tt <u>c</u> gt <u>c</u> ca <u>g</u> gag aac gr <u>g</u> aac gac gg <u>c</u> gcc gtc acc gac gg <u>c</u> ct <u>c</u> gac gt <u>g</u>											63% 35%
G L F V O D N V A D N K V T D R L A S V											

SUBSTITUTE SHEET (RULE 26)

Fig. 1 Sheet 1

2/6

362/121	gag	act	gat	atg	cct	cta	aca	tta	aga	cag	cat	aga	ata	agt	gcg	aca	tca	tca	tcg	gaa	58%
gag	acc	gac	atg	ccc	ctc	acc	ctg	cgc	cag	cac	cgc	atc	agc	gcg	acc	tcc	tcc	tcg	gag	30%	
E	T	D	M	P	L	T	L	R	Q	H	R	I	S	A	T	S	S	S	E		
422/141	gag	agt	aac	aaa	ggt	caa	aga	cag	ttt	act	gtt	act	gtt	tcg	atcg	gcc	ccc	ccg	acc	gtc	47%
gag	agc	agc	aac	aag	ggc	cag	cgc	cag	ttt	acc	gtc	tcg	atcg	gcc	ccc	ccg	acc	gac	gac	gtc	30%
E	S-	N	K	G	Q	R	Q	L	T	V	S	T	A	P	P	T	D	V			
482/161	agc	ctg	ggg	gac	gag	<u>ctc</u>	cac	tta	gac	ggc	gag	gac	gtt	gctg	atg	gctg	cat	gcc	gac	gctg	47%
agc	ctg	ggg	gac	gag	<u>ctc</u>	cac	tta	gac	ggc	gag	gac	gtt	gctg	atg	gctg	cat	gcc	gac	gctg	30%	
S	L	G	D	E	L	H	L	D	G	E	D	V	A	M	A	H	A	D	A		
542/181	cta	gac	gat	ttc	gat	ctg	gac	atg	ttt	ggg	gac	ggg	gat	tcc	ccg	ggg	ccg	ggg	ttt	acc	40%
cta	gac	gat	ttc	gat	ctg	gac	atg	ttt	ggg	gac	ggg	gat	tcc	ccg	ggg	ccg	ggg	ttt	acc	40%	
L	D	F	D	L	D	M	L	G	D	G	D	S	P	G	P	G	F	T			
602/201	ccc	cac	gac	tcc	gcc	ccc	tac	ggc	gct	ctg	gat	atg	ggg	ttt	gag	ttt	gag	ttt	acc		
ccc	cac	gac	tcc	gcc	ccc	ccc	tac	ggc	gct	ctg	gat	atg	ggg	ttt	gag	ttt	gag	ttt	acc		
P	H	D	S	A	P	Y	G	A	L	D	M	A	D	F	E	F	E	Q	M		
662/221	ttt	acc	gat	gcc	cct	gga	att	gac	gag	tac	ggt	ggg	tag	Bgl	II						
ttt	acc	gat	gcc	cct	gga	att	gac	gag	tac	ggt	ggg	tag									
F	T	D	A	L	G	I	D	E	Y	G	G	*									

Fig. 1 Sheet 2

3/6

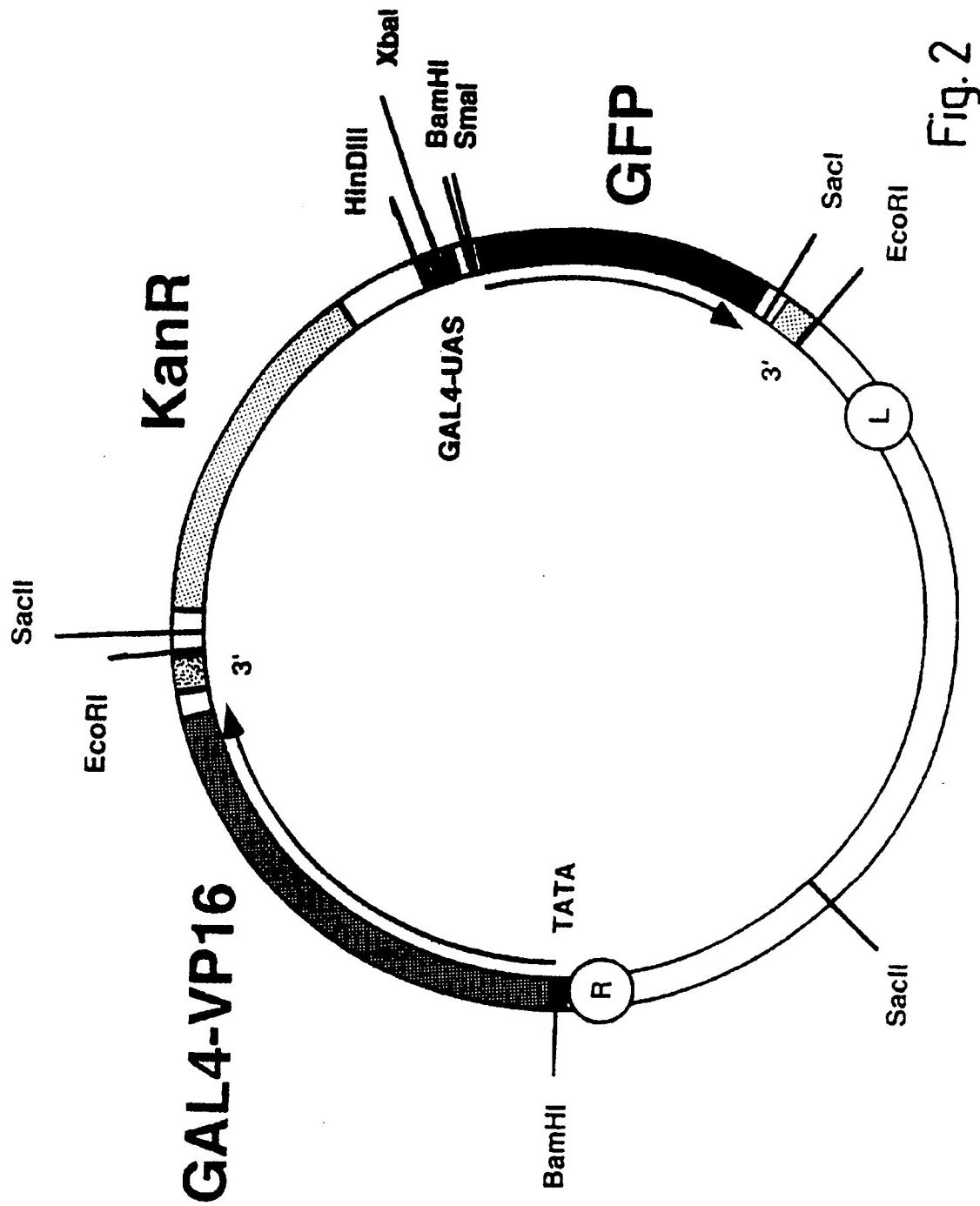


Fig. 2

4/6

Modified Ti plasmid right border

TAGGTITACCGCCAATATATCCTGTCAAACACT GGATCTTCGCAAGACCCTT
|-----25 bp repeat-----|
mRNA start
|
CCTCTATATAAGGAAGTTCATTCATTGGAGAGGACA.. GAL4 gene..
TATA Box

Fig. 3

5/6

X
UAS-nGUS



Fig. 4a

Fig. 4c

X-gluc stained
brightfield

6/6



Fig. 4 b



Fig. 4d

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 97/00406

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/82 A01H5/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N A01H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	NATURE, vol. 334, 18 August 1988, pages 631-633, XP000645136 MA J ET AL: "YEAST ACTIVATORS STIMULATE PLANT GENE EXPRESSION" cited in the application see the whole document ---	1-21
A	DEVELOPMENT, vol. 118, no. 2, 1993, pages 401-415, XP000674470 BRAND A. AND PERRIMON N.: "Targeted gene expression as a means of altering cell fates and generating dominant phenotypes" cited in the application * see the whole document, esp. p.413,col.1,par.3 *	1-21 -/-

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- *&* document member of the same patent family

2

Date of the actual completion of the international search

6 June 1997

Date of mailing of the international search report

13.06.97

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+ 31-70) 340-3016

Authorized officer

Kania, T

INTERNATIONAL SEARCH REPORT

Int. Application No
PCT/GB 97/00406

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	NATURE, vol. 362, 15 April 1993, pages 630-632, XP002032417 GREIG S. AND AKAM M.: "Homeotic genes autonomously specify one aspect of pattern in the Drosophila mesoderm" see the whole document ---	1-21
A	PLANT CELL REPORTS, vol. 14, 1995, pages 773-776, XP000650425 REICHEL C. ET AL.: "Inefficient expression of the DNA-binding domain of Gal-4 in transgenic plants" cited in the application see page 776 ---	1
A	WO 90 01551 A (ROGERS JOHN C) 22 February 1990 * see esp. ex. 4+5 *	1
A	WO 95 33055 A (CALGENE INC ;LASSNER MICHAEL (US); METZ JAMES GEORGE (US)) 7 December 1995 see page 12, line 11-32 ---	1
A	EP 0 359 472 A (LUBRIZOL GENETICS INC) 21 March 1990 see page 3, line 52 - page 4, line 10 ---	1
A	EP 0 385 962 A (MONSANTO CO) 5 September 1990 see page 2 - page 5 ---	1
A	EP 0 589 841 A (CIBA GEIGY AG) 30 March 1994 see example 16 ---	17
A	GENES AND DEVELOPMENT, vol. 9, no. 14, 15 July 1995, pages 1797-1810, XP000674520 SUNDARESAN V. ET AL.: "Patterns of gene action in plant development revealed by enhancer trap and gene trap transposable elements" cited in the application see the whole document ---	1-21
T	CA 2 150 039 A (UNIVERSITY OF WARWICK) 9 August 1996 see the whole document -----	1-21

2

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int. Appl. No.

PCT/GB 97/00406

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9001551 A	22-02-90	AU 638409 B AU 4037289 A EP 0428572 A JP 4500153 T	01-07-93 05-03-90 29-05-91 16-01-92
WO 9533055 A	07-12-95	CA 2168042 A CN 1129014 A EP 0711351 A JP 9501325 T	07-12-95 14-08-96 15-05-96 10-02-97
EP 0359472 A	21-03-90	AT 132193 T AU 623429 B AU 4118289 A CN 1044298 A DE 68925253 D DE 68925253 T EP 0682115 A ES 2083384 T JP 2186989 A US 5380831 A US 5567600 A US 5567862 A	15-01-96 14-05-92 15-03-90 01-08-90 08-02-96 13-06-96 15-11-95 16-04-96 23-07-90 10-01-95 22-10-96 22-10-96
EP 0385962 A	05-09-90	AU 638438 B AU 5163090 A CA 2024811 A EP 0413019 A IL 93513 A JP 3504333 T WO 9010076 A US 5500365 A	01-07-93 26-09-90 25-08-90 20-02-91 14-11-95 26-09-91 07-09-90 19-03-96
EP 0589841 A	30-03-94	US 5409823 A BR 9303877 A CA 2106718 A HU 69965 A	25-04-95 29-03-94 25-03-94 28-09-95
CA 2150039 A	09-08-96	NONE	



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ :		A1	(11) International Publication Number: WO 96/27675
C12N 15/82, 15/12, C07K 14/435, G01N 33/52, 33/53			(43) International Publication Date: 12 September 1996 (12.09.96)
(21) International Application Number: PCT/GB96/00481		(81) Designated States: AU, CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(22) International Filing Date: 4 March 1996 (04.03.96)			
(30) Priority Data: 9504446.7 6 March 1995 (06.03.95) GB		Published <i>With international search report.</i>	
(71) Applicant (for all designated States except US): MEDICAL RESEARCH COUNCIL [GB/GB]; 20 Park Crescent, London WIN 4AL (GB).			
(72) Inventors; and (75) Inventors/Applicants (for US only): HASELOFF, James, Phillip [AU/AU]; Glebe Cottage, Church Lane, Comberton, Cambridgeshire CB3 7ED (GB). HODGE, Sarah [GB/GB]; 15 St Luke's Street, Cambridge CB4 3DA (GB). PRASHER, Douglas [US/US]; 72 Goletta Drive, East Falmouth, MA 02536-3964 (US). SIEMERING, Kirby [AU/AU]; 8 Portugal Place, Cambridge CB5 8AF (GB).			
(74) Agent: KEITH W. NASH & CO.; 90-92 Regent Street, Cambridge CB2 1DP (GB).			
(54) Title: JELLYFISH GREEN FLUORESCENT PROTEIN (GFP) EXPRESSION IN PLANTS			
<pre> 201/61 231/71 NdeI gtc act act ttc tct tat ggt gtt caa tgc ttt tca aga tac cca gat [REDACTED] aas cgg gtc act act ttc tct tat ggt gtt caa tgc ttt tca aga tac cca gat [REDACTED] aag cgg V T F S Y G V Q C F S R Y P D H M K R </pre>			
<pre> 261/81 291/91 cat gac ttt ttc aag agt gcc atg ccc gaa ggt tat gta cag gaa aga act ata ttt ttc cac gac ttc ttc aag agc gcc atg cot gag gga taa ggt cag gag egg aac atc ttc ttc H D F F K S A H P B G Y V Q E R T I F F </pre>			
<pre> 321/101 351/111 ↓ aaa gat gac ggg aac tac aag aca cgt gct gaa gtc aag ttt gaa ggt cat acc ctt atc aag gac gac ggg aac tac aag aca cgt gct gaa gtc aag ttt gaa ggg gaa acc ctc gtc K D D G M Y K T R A E V K F X G D T L V </pre>			
<pre> 381/121 411/131 aat aca atc gag tta aca gat att gat ttt aca gaa cat cca aac att ctt cca cac aac aac agg atc gag ott aag gga atc gat ttc aag gag gac gga aac atc ctc ggc cac aag M R I E L X G I D F K E D G N I L G H K </pre>			
<pre> 441/141 Acc/ 471/151 ↓ ttt gaa tac sac tat aac tca cac sat [REDACTED] atc atg gca gac aaa caa aag aat gga ttt gaa tac sac tat aac too cac aac [REDACTED] atc atg gca gac aaa caa aag aat gga L E Y W Y N S H H V Y I M A D X Q K N G </pre>			
<p>top sequence = <i>Aequorea victoria GFP</i>, lower sequence = <i>mGFP4</i> Intron sequences are underlined and the cryptic splice junctions are arrowed. Mutated nucleotides are shown outlined. Nucleotide and amino acid numbering starts at the initiation codon.</p>			
(57) Abstract			
<p>Disclosed is a DNA sequence encoding Green Fluorescent Protein (GFP), the sequence being modified relative to the wild type sequence so as to allow for more efficient expression in a plant cell of a functional GFP polypeptide. Also disclosed is a modified GFP polypeptide comprising an amino acid substitution, relative to the wild type protein sequence, which exhibits useful characteristics when expressed in many different types of host cell.</p>			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgyzstan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	KZ	Kazakhstan	SG	Singapore
CH	Switzerland	LI	Liechtenstein	SI	Slovenia
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovakia
CM	Cameroon	LR	Liberia	SN	Senegal
CN	China	LT	Lithuania	SZ	Swaziland
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	LV	Latvia	TG	Togo
DE	Germany	MC	Monaco	TJ	Tajikistan
DK	Denmark	MD	Republic of Moldova	TT	Trinidad and Tobago
EE	Estonia	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	UG	Uganda
FI	Finland	MN	Mongolia	US	United States of America
FR	France	MR	Mauritania	UZ	Uzbekistan
GA	Gabon			VN	Viet Nam

JELLYFISH GREEN FLUORESCENT PROTEIN (GFP) EXPRESSION IN PLANTS

Field of the Invention

This invention relates to improvements in gene expression, especially improvements in expression of the Green Fluorescent Protein (GFP) gene, and to a method of detecting the presence and/or expression in a host of a gene of interest.

Background of the Invention

Genes encoding β -glucuronidase and β -galactosidase have been used as reporters for gene expression in plants. Using these reporter genes, transformed tissues or patterns of gene expression can be identified histochemically, but this is generally a destructive test and is not suitable for assaying primary transformants, nor for following the time course of gene expression in living plants, nor as a means of rapidly screening segregating populations of seedlings. There is thus a general need for improved reporters of gene expression, but especially reporter genes for use in plants. Candidates might be found among proteins having intrinsic fluorescence.

Proteins with high intrinsic fluorescence are involved in photosynthesis and bioluminescence, and in most cases possess a protein-bound chromophore. For example, the highly fluorescent phycobiliproteins require complex tetrapyrrole groups, and the blue and yellow fluorescent proteins from *Vibrio fischeri* must bind lumazine and flavin mononucleotide, respectively. This requirement for an external chromophore complicates the use of these proteins as reporters for gene expression. However, the green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* does not share this requirement for an external chromophore.

Aequorea victoria are brightly luminescent, with light appearing as glowing points around the margin of the jellyfish umbrella. Light arises from yellow tissue masses which each consist of about 6000-7000 photogenic cells (Davenport & Nichol, 1955 Proc. Roy. Soc. Ser. B 144, 399-411). The cytoplasm of these cells is densely packed with fine granules

of about 0.2 μ m diameter which are enclosed by a unit membrane and contain the components necessary for bioluminescence (Anderson & Cormier, 1973 *J. Biol. Chem.* 248, 2937-2943). The components include a Ca⁺⁺ activated photoprotein, aequorin, that emits blue-green light, and an accessory green fluorescent protein (GFP) which accepts energy from aequorin and re-emits it as green light.

GFP is an extremely stable protein of 238 amino acids. The fluorescent properties of the protein are unaffected by prolonged treatment with 6M guanidine HCl, 8M urea or 1% SDS, and two day treatment with various proteases such as trypsin, chymotrypsin, papain, subtilisin, thermolysin and pancreatin at concentrations up to 1 mg/ml fail to alter the intensity of GFP fluorescence. GFP is stable in neutral buffers up to 65°C, and displays a broad range of pH stability from 5.5 to 12 (Bokman & Ward, 1981 *Biochem. Biophys. Res. Comm.* 101, 1372-1380). The protein is intensely fluorescent, with a quantum efficiency of approximately 80% and molar extinction coefficient of around 4.5x10⁴. GFP absorbs light maximally at 395 nm and has a smaller absorbance peak at 475nm, and fluorescence emission peaks at 509nm, with a shoulder at 540nm (Morise *et al.*, 1974 *Biochemistry* 13, 2656-2662). Researchers have successfully cloned and sequenced both the cDNA and genomic DNA sequences coding for *A. victoria* GFP (Prasher *et al.*, 1992 *Gene* 111, 229-233).

The fluorescence of GFP has been well characterised (Inouye & Tsuji, 1994 *FEBS Letters* 33817, 277-280 and *FEBS Letters* 34472, 211-214) and appears to be due to a unique covalently-attached chromophore which is formed post-translationally by cyclisation and oxidation of the residues 65-67 (Ser-Tyr-Gly) within the protein (Cody *et al.*, 1993 *Biochemistry* 32, 1212-1218; Heim *et al.*, 1994 *PNAS* 91, 12501-12504). Several genomic and cDNA clones of *gfp* have been obtained from a population of *A. victoria*. The *gfp* gene contains at least three introns, and the sequences derived from the cDNA have been used for protein expression studies in *Escherichia coli*, *Caenorhabditis elegans* (Chalfie *et al.*, 1994 *Science* 263, 802 *et seq.*) and *Drosophila melanogaster* (Wang & Hazelrigg, 1994 *Nature* 369, 400 *et seq.*). Fluorescent protein was produced in these different cell types and there appears to be little requirement for specific additional factors for post-translational modification of the protein, which may be

autocatalytic or require common factors. Recently, modified forms of GFP have also been made and studied, which forms have rather different fluorescence characteristics compared to the wild type (Heim *et al.*, 1994 PNAS 91, 12501-12504; Delagrave *et al.*, 1995 Bio/Technology 13, 151 *et seq.*; and Heim *et al.*, 1995 Nature 373, 663-664).

Although GFP has some advantages as a fluorescent reporter molecule, expression has been reported to be problematic in some experimental systems (Cubitt *et al.*, 1995 Trends Biochem. Sci. 20, 448-455). Expression of GFP in mammalian cells has been described as highly variable (Rizutto *et al.*, 1995 PNAS 92, 11899-11903); Kaether & Gerdes 1995, FEBS Lett. 369, 267-271; Pines 1995, Trends Genet. 11, 326-327) often requiring a strong promoter and decreased incubation temperature for good results (Ogawa *et al.*, 1995 PNAS 92, 11899-11903). Other researchers have found that a lower incubation temperature also favours the development of fluorescence during expression of GFP in bacteria (Heim *et al.*, 1994 PNAS 91, 12501-12504; Webb *et al.*, 1995 J. Bact. 177, 5906-5911) and yeast (Lim *et al.*, 1995 J. Biochem. 118, 13-17). In yeast, this phenomenon has been attributed primarily to more efficient maturation of GFP to the fluorescent form at lower temperatures. The present inventors have sought to express modified forms of GFP in various hosts.

Summary of the Invention

In a first aspect the invention provides a DNA sequence encoding Green Fluorescent Protein (GFP), the sequence being modified relative to the wild type sequence so as to allow for more efficient expression in a plant cell of a functional GFP polypeptide. Preferably the modified sequence is capable of efficient expression in a dicotyledonous plant, such as *Arabidopsis*.

The term "GFP" as used herein is intended to refer to a polypeptide possessing many of the properties of the naturally occurring protein, and particularly exhibiting intrinsic fluorescence. As will be apparent to those skilled in the art, the polypeptide need not necessarily fluoresce in the "green" part of the visible spectrum, as the fluorescence properties of the polypeptide (including the wavelength of fluorescence) may be

substantially altered by one or more mutations.

The GFP polypeptide will generally have substantially the amino acid sequence of the wild type protein (as disclosed by Prasher *et al.*), but will preferably comprise one or more amino acid differences defined, and discussed in greater detail, below.

The GFP-coding DNA sequence is advantageously modified so as to reduce the probability of an RNA sequence transcribed therefrom being subject to erroneous splicing in a plant cell. The DNA sequence is conveniently modified so as to comprise a plurality of nucleotide substitutions relative to the wild type sequence, which substitutions serve to reduce, or preferably entirely prevent, excision from the transcribed RNA of the portion corresponding to nucleotides 400-483 of the DNA sequence. It has surprisingly been found by the present inventors that this portion of the sequence tends to be recognised in plant cells (particularly dicotyledenous plants) as an intron, which is therefore excised by splicing of the RNA.

Preferably the nucleotide substitutions in the DNA sequence serve to decrease the A/U% content of the transcribed RNA, which is believed to decrease the likelihood of the sequence being treated by a plant cell as representing an intron. Desirably the substitutions particularly decrease the A/U% content of the region corresponding to nucleotides 400-483. Conveniently the nucleotide substitutions are such as to preserve the amino acid sequence of the encoded polypeptide substantially unchanged in the portion encoded by nucleotides 400-483. Other substitutions may advantageously be made to decrease the similarity between the GFP RNA sequence and the plant intron recognition consensus sequence (see Figure 2).

In addition to nucleotide substitutions in the portion 400-483 of the DNA sequence, a number of other modifications may advantageously be made. For example, substitutions may be made downstream (i.e. 3') and/or upstream (i.e. 5') of nucleotides 400-483, which substitutions will conveniently serve to further reduce the A/U content of the transcribed RNA.

The invention also includes within its scope RNA sequences capable of being transcribed from the modified DNA sequence (i.e. RNA sequences transcribed from the modified DNA sequence, or an RNA having a sequence such that it could be synthesised by transcription from the modified DNA sequence).

Advantageously a number of other modifications, in addition to those specified above, may also be made to the GFP-coding sequence. For example, the sequence will typically further comprise transcription and translation signals (e.g. promoters, enhancers) and/or localisation signals recognised in plants. Localisation signals may direct the expressed polypeptide to: the nucleus (e.g. SV40 large T antigen localisation signal, particularly in combination with other polypeptide sequences, which have been found to increase the efficiency of the signalling); mitochondria (e.g. cytochrome C oxidase subunit IV); endoplasmic reticulum - ER (e.g. the signal sequence from carboxypeptidase Y, or that from *Arabidopsis* basic chitinase); or microbodies (e.g. peroxisomes). It may even be possible to target the modified GFP polypeptide to the plant cell wall. Localisation of GFP may be highly desirable when expression occurs at high levels, so as to minimise possible toxicity to host cells.

The sequence may advantageously be further modified in accordance with the manner described in the prior art (e.g. as disclosed by Heim *et al.*, 1994, 1995, or Delagrange *et al.*, 1995, as cited previously). Whilst in general the DNA sequence is modified in such a way as to preserve the wild type amino acid sequence, it has been found that amino acid changes at specific residues are in fact desirable. In particular, the sequence may be modified so as to comprise an amino acid substitution at one or both of amino acid residues 163 and 175. Changes at these positions are found to alter the characteristics of the polypeptide in an unexpected and favourable manner.

In particular, amino acid substitutions at residue 163 and/or 175 (valine and serine respectively, in the wild type protein) have favourable effects on the characteristics of the GFP polypeptide when expressed in many different host cells (e.g. bacterial, yeast etc.), and such substitutions may advantageously be included independently of any modification of the DNA sequence made for increased efficiency of expression in plants.

In a second aspect therefore the invention provides a modified GFP polypeptide comprising an amino acid substitution relative to the wild type protein at residue 163 and/or 175. Substitution at either residue, in isolation, has surprisingly been found to increase the thermotolerance of the polypeptide. Maximal thermotolerance is obtained by causing substitution at both residues.

Advantageously valine 163 is substituted by alanine, or by a related amino acid (i.e. those having an aliphatic side chain: glycine, leucine and isoleucine; or those having an aliphatic hydroxyl side chain: serine and threonine).

Advantageously serine 175 is substituted by glycine, or by a related amino acid (i.e. those having an aliphatic side chain: alanine, leucine and isoleucine; or those having an aliphatic hydroxyl side chain: serine and threonine).

It is preferred that the modified GFP polypeptide comprises substitutions at both residues, conveniently 163→alanine and 175→glycine.

The modified GFP may additionally comprise other sequence differences relative to the wild type protein, particularly in, or immediately adjacent to, residues 65-67 (which residues give rise to the chromophore).

Such a nucleic acid sequence is useful for example, as a marker, or as a reporter gene, in a wide variety of host cells (e.g. mammalian, bacterial, fungal, yeast or plant cells). Conveniently, the nucleic acid sequence may be further modified for expression in a particular host cell. For example, where the thermotolerant GFP-coding sequence is to be expressed in a plant cell it will conveniently be modified in accordance with the first aspect of the invention.

In a third aspect, the invention provides a nucleic acid construct comprising a nucleic acid sequence in accordance with the first aspect of the invention. In particular the construct is preferably an expression vector, comprising one or more regulatory signals (such as promoters etc.) and is preferably suitable for use in a plant cell. The construct will

desirably include one or more restriction endonuclease sites, suitable for the insertion into the construct of other nucleic acid sequences, which in a preferred embodiment may be inserted in frame with the sequence of the invention.

The invention also provides a host cell, conveniently a plant cell, into which has been introduced a sequence in accordance with the first aspect of the invention.

Processes for introducing DNA into plant cells (typically by transformation) are not 100% efficient. Accordingly, it is generally desirable for the DNA introduced into the plant cell to confer one or more distinctive characteristics upon the plant cell, which characteristic(s) serve to mark those cells which have taken up the DNA. The fluorescent properties of GFP constitute such a distinctive characteristic ("marker"). In a preferred embodiment the invention thus provides a plant cell transformation vector comprising the sequence of the invention. Further, the invention provides a method of screening plant cells, comprising introducing into at least some of a plurality of plant cells a DNA construct comprising a sequence in accordance with the invention, maintaining the cells under suitable conditions for an appropriate length of time so as to allow expression of a modified GFP from the construct, and selecting those cells which exhibit GFP-mediated fluorescence. "Suitable conditions" and "an appropriate length of time" are well known to those skilled in the art from standard texts.

In a preferred embodiment, the vector further comprises a sequence of interest which, preferably, is present in frame with the modified GFP-coding sequence.

In a fourth aspect the invention thus provides a method of detecting the expression in a plant of a sequence of interest, comprising causing the sequence of interest to be present in frame with a modified GFP-coding sequence in accordance with the first aspect of the invention so as to form a modified GFP/sequence of interest fusion, introducing the fusion into a plant, and monitoring the fluorescence thereof. GFP-mediated fluorescence is thus an indicator of expression of the sequence of interest.

In yet another aspect, the invention provides a nucleic acid construct comprising a

sequence in accordance with the second aspect of the invention. The nucleic acid construct will desirably have many of the features of the nucleic acid construct in accordance with the third aspect of the invention. It will be apparent however that the construct may be useful in many different types of host cell, and may be constructed accordingly.

The invention will now be further described by way of illustration and with reference to the accompanying drawings, in which:

Figure 1A shows the sequences introduced, via PCR, flanking the GFP-coding sequence;

Figure 1B is a confocal micrograph of transformed yeast cells expressing GFP;

Figure 2A is a photograph showing agarose gel electrophoresis analysis of PCR products;

Figure 2B is a schematic illustration of the portion of DNA not represented in the mis-spliced mRNA produced in plants from the wild type GFP-coding sequence;

Figure 3A is a photograph showing the DNA sequence determination of the reverse transcript produced from mis-spliced mRNA;

Figure 3B is a comparison between a portion of the GFP wild type sequence and a plant intron consensus sequence;

Figure 4 shows a comparison of part of the wild type *A. victoria* GFP sequence with a modified GFP-coding sequence in accordance with the invention;

Figure 5 is a series of confocal micrographs (at different magnification) showing parts of a plant expressing a modified GFP-coding sequence in accordance with the invention;

Figure 6 shows a comparison of part of three modified GFP-coding sequences in accordance with the invention, together with the amino acid sequences encoded thereby;

Figure 7 is a graph of relative fluorescence (arbitrary units) against time (minutes) for *E. coli* strains expressing modified GFP (open squares) or modified, mutated GFP (filled circles);

Figure 8 is a photograph of a Western blot, probed with anti-GFP antibody;

Figure 9 is a bar chart showing the amount of fluorescence associated with cultures expressing modified GFP (open columns) or modified, mutated GFP (shaded columns) incubated at four different temperatures;

Figure 10 is a graph of fluorescence against time (minutes) for yeast cultures at 25°C or 37°C, the cultures having been grown initially in anaerobic conditions, with oxygen introduced at time zero;

Figure 11 is a picture of a Western blot showing expression of modified GFP or modified, mutated GFP (GFPA) by *E. coli* cultures at 25 or 37°C, with comparison between soluble and insoluble culture fractions;

Figure 12 is a graph of absorbance against wavelength for soluble (filled circles) or insoluble (open circles) GFP;

Figure 13 is a picture of yeast cultures, grown at 25 or 37°C and expressing modified GFP, or modified, mutated GFP (GFPA);

Figure 14 is a graph of fluorescence against wavelength (nm), showing the excitation spectra (squares) and emission spectra (circles) respectively, of modified GFP (solid lines) and two mutated forms of modified GFP, GFPA (dashed lines) and GFP5 (dotted lines);

Figure 15 is a comparison of the nucleotide sequence of wild-type *gfp* and a modified gene *m-gfp5*, and the polypeptides encoded thereby. Nucleotide sequence differences are shown in bold. The m-GFP5 amino acid sequence is shown beneath the nucleotide

sequence. The three amino acid differences between the encoded polypeptides are indicated:

Figure 16 shows the sequence of another modified *m-gfp* gene, termed *m-gfp5-ER*, and the amino acid sequence of the polypeptide encoded thereby; and

Figure 17 shows a number of confocal micrographs (A-H) of *Arabidopsis* seedlings expressing modified *gfp* genes in accordance with the invention.

Examples

Example 1

Construction of a *gfp* expression cassette

A synthetic *gfp* gene was constructed using the polymerase chain reaction (PCR). The plasmid pGFP10.1 (described by Prasher *et al.*, 1992 Gene 111, cited above) contains a cloned *A. victoria* *gfp* cDNA, and was used as template for PCR amplification (with *Thermococcus litoralis* Vent polymerase) with synthetic oligomer primers which were used to incorporate new sequences flanking the GFP coding sequence.

The sequence of the primer oligonucleotides was:

GGCGGATCCAAGGAGATATAACAATGAGTAAAGGAGAAGAACTTTCACT (Seq. ID No. 1) and GGCGAGCTCTATTGTATAGTTCATCCATGCC (Seq. ID No. 2).

The newly-incorporated sequences are shown in Figure 1A. Referring to Figure 1A, the sequence existing in pGFP10.1 is shown italicised. The added sequences are shown in normal type. These included: recognition sites for the restriction endonucleases *Bam*HI and *Sac*I placed at the 5' and 3' termini of the amplified fragment; a Shine-Delgarno ribosome binding site (RBS) sequence positioned upstream of the initiation codon to ensure efficient translation of the transcribed gene in *E. coli*, and the sequence AACAA inserted between positions -4 and -1 for efficient translation in plants.

The PCR-amplified fragment was subcloned into pUC119 for bacterial expression, and

into an episomal yeast plasmid vector, pVT103-U (Vernet *et al.*, 1987 Gene 52, 225-233) which contains a yeast 2 μ M origin of replication and a truncated form of the yeast ADH1 promoter to allow high level expression of the cloned GFP gene in *Saccharomyces cerevisiae*. *S. cerevisiae* MGLD-4a (a. *leu2*, *ura3*, *his3*, *trp1*, *lys2*) cells were transformed using the lithium acetate method described by Ito *et al.* (1983).

Transformed *E. coli* and yeast bearing the recombinant plasmid were observed to produce brilliantly fluorescent colonies under long wavelength UV illumination using a hand-held lamp. Interestingly, a high degree of sectoring was seen in yeast colonies containing the episomal form of the *gfp* cDNA; the sectoring was eliminated by integration of the *gfp* cDNA at the yeast URA3 locus and presumably reflects the instability of the 2 μ M-based episome and/or some toxic effects of GFP expression. This observation also indicated the utility of GFP as a simple cell-autonomous marker. Examination of transformed yeast cells by confocal microscopy showed that the protein was predominantly distributed throughout the cytoplasm (Fig. 1B).

After the PCR amplified *gfp* cDNA was shown to correctly produce fluorescent protein product in yeast, the sequence was inserted between the *Bam*H I and *Sac*I sites in the plant transformation vector pBI121 behind the 35S promoter (Jefferson *et al.*, 1987). *A. tumefaciens* strain LBA4044 was transformed with the *gfp*-containing plasmid by electroporation. Roots of *Arabidopsis thaliana* C24 were transformed using the protocol of Valvekans *et al.* (1988). Transgenic callus and shoots were screened for GFP expression using an inverted fluorescence microscope (Leitz DM-IL) fitted with an appropriate filter set (Leitz-D). However, at no stage during the transformation procedure was GFP-related fluorescence detected by UV lamp illumination, or by epifluorescence microscopy. This lack of fluorescence was unexpected and surprising in view of the fluorescence exhibited previously by the transformed bacteria and yeast cells.

gfp is mis-spliced in *Arabidopsis*

The successful expression of GFP in *Arabidopsis* requires proper production of the apoprotein, before post-translational modification to form the chromophore. The inventors therefore used PCR-based methods to verify the correct insertion of the 35S

promoter-driven *gfp* cDNA, and to check mRNA transcription and processing in transformed plantlets. Nucleic acids were extracted from plantlets and either treated with RNase, or DNase treated and reverse transcribed using oligo(dT)₈ primer. The *gfp* sequences in these extracts were therefore derived from genomic DNA or transcribed mRNAs, respectively. The *gfp* sequence was PCR-amplified from these separate extracts and products were analysed by restriction endonuclease digestion, as shown in Figure 2A.

In Figure 2A, the RNase-treated DNA sample ("mRNA") is shown on the left of each pair of samples, whilst the DNase-treated/reverse transcribed sample is shown on the right of each pair. The samples were either loaded onto gel without prior restriction ("uncut", extreme left hand pair of samples) or loaded after prior digestion with (from left to right): *NcoI*; *RsaI*; *DraI*; *AccI*; *HincII*; or *AvaII*. It can be seen that whilst the expected product was obtained after amplification of the gene, RT-PCR of mRNA sequences gave rise to a truncated product. This product was 80-90 base pairs shorter than expected and was uncut by the restriction endonucleases *DraI* and *AccI*, whilst the gene sequences contained unique recognition sites for these enzymes. The inventors established that this is consistent with a small deletion within the *gfp* coding sequence as shown in Figure 2B. In this figure, the shaded portion represents that missing from the mRNA-derived RT-PCR amplified sequences.

The shortened RT-PCR product was cloned and sequenced (Fig. 3A), and a deletion of 84 nucleotides between residues 400-483 was located. The nucleotide sequences bordering this deletion are shown in figure 3B, and demonstrate similarity to known plant introns. The sequence across the splice site (marked with an arrow in Figure 3A) thus reads (5' to 3') ...AG/AC.... . Matches were found for important residues at the 5' and 3' splice sites (reviewed by Luehrs *et al.*, 1994) and the excised *gfp* sequence contains a high predicted A:U content (68%) which has also been shown to be important for recognition of plant introns (Wiebauer *et al.*, 1988; Hanley & Schuler, 1988; Goodall & Filipowitz, 1989, 1991). It is therefore likely that this 84 nucleotide region of the jellyfish *gfp* cDNA sequence is recognised as an intron when transcribed in *Arabidopsis*, resulting in the production of a defective protein product. It should be noted that the borders of this cryptic intron do not coincide with any of the natural spliced junctions

found after processing of the *gfp* mRNA in *A. victoria*.

Modification of the *gfp* gene

The jellyfish gene was mutated to produce a modified *gfp* (*m-gfp*) suitable for expression in *Arabidopsis*, as described below.

Two mutagenic oligonucleotides were synthesized, a 122-mer:

GATCATATGAAGCGGCACGACTTCTCAAGAGGCCATGCCTGAGGGATACG
TGCAGGAGAGGACCATCTTCTCAAGGACGACGGAACTACAAGACACGTG
CTGAAGTCAAGTTGAGGG (Seq. ID No. 3),

and a 126-mer:

GATGTATACTGTTGGGAGTTGTAGTTGTATTCCAACTTGTGGCCGAGGATG
TTTCCGTCTCCTTGAAATCGATTCCCTTAAGCTCGATCCTGTTGACGAGGGT
GTCTCCCTCAAAC TTGACTTC (Seq. ID No. 4).

The oligonucleotides were purified by electrophoresis in a 5% polyacrylamide gel containing TBE and 7M urea. The gel was stained briefly with 0.05% toluidine blue, and the full-length oligonucleotides were excised, and eluted overnight in 0.5M ammonium acetate, 0.1mM Na₂EDTA, 0.1% SDS. The oligonucleotides share 17 nucleotides of complementarity at their 3' termini, and were annealed and elongated after several rounds of thermal cycling with Vent polymerase. The extended product was cloned between the *Nde* I and *Acc* I sites of *gfp*. The mutant clones were screened for the presence of the diagnostic restriction endonuclease sites, *Cla* I, *Ava* II, and the desired fragment (*m-gfp*) was subcloned into M13 and its sequence verified by DNA sequencing using the dideoxynucleotide chain termination technique with T7 DNA polymerase.

The modifications introduced by the synthetic oligonucleotides were intended to alter the sequences which might be involved in 5' splice site recognition and to decrease the A:U content of the putative intron, as shown in Figure 4. In the Figure, the upper DNA sequence is that of a portion of the wild type *A. victoria* GFP. The lower DNA sequence

is that of a portion of a modified GFP-coding sequence. The corresponding amino acid sequence is shown beneath the DNA sequences. Modified nucleotides are shown outlined. All DNA modifications affect only codon usage, and the *m-gfp*-encoded amino acid sequence is identical to that of the wild-type jellyfish polypeptide. The "pseudo-intron" sequence is underlined and the cryptic splice junctions are arrowed. Nucleotide and amino acid residue numbering (on the left and right, respectively, of the oblique stroke) start from the initiation codon. The boxed hexanucleotide sequences are *NdeI* and *AccI* recognition sites respectively.

The *m-gfp* sequence was inserted behind the 35S promoter in pBI121, and introduced into *Arabidopsis* using the root transformation technique. Brightly green fluorescent cells were seen after co-cultivation with *Agrobacterium*. As shoot regeneration progressed, explants with different levels of green fluorescence could be observed. Regenerating callus and shoots develop a bright red autofluorescence due to the formation of chlorophyll within the tissues, and with the brightest *m-gfp* transformants the green fluorescence was clearly detectable against this autofluorescent background using a hand held UV lamp. This was similar to the levels of green fluorescence seen in transformed yeast and *E. coli*. However, these very bright *Arabidopsis* transformants regenerated and set seed rather poorly. Nevertheless, seeds were obtained from over 50 transformed lines, allowed to germinate, and screened by epifluorescence microscopy. Several of the brightest lines were used for confocal laser scanning microscopy.

Confocal microscopy of living plants

The fluorescence properties of GFP and chlorophyll allow the use of fluorescence microscopes equipped with common filter sets for fluorescein and rhodamine for dual imaging in plant cells. Intact five day old *m-gfp*-transformed *Arabidopsis* seedlings were mounted in water for confocal laser scanning microscopy. GFP fluorescence could be clearly visualised in the transformed tissues, and chloroplasts provided a very effective counter fluor in the upper parts of the plant. Optical sectioning of the *m-gfp* transformed plants gives selective visual access to the internal details of living plant structure, as shown in Figure 5, without any need for staining or dissection. For example, median longitudinal sections of root tips can be simply obtained by adjusting the microscope

depth of focus, and confocal imaging allows the resolution of subcellular details. GFP is found throughout the cytoplasm, but appears to accumulate within the nucleoplasm. It appears excluded from vacuoles, organelles and other bodies in the cytoplasm, and is excluded from the nucleolus. Similarly, in optical sections of cotyledon and hypocotyl tissues, GFP is found throughout the cytoplasm and nucleoplasm. The relationship of cells within the tissues is clearly discernible. In highly vacuolate epidermal cells in the root and hypocotyl, GFP fluorescence allows visualisation of trans-vacuolate cytoplasmic threads, and the thin cytoplasmic strands which underly the cell wall and which may be aligned with cytoskeletal elements. The movement of organelles through cytoplasmic streaming could also be observed in these living cells.

Example 2

Isolation and characterisation of a bright mutant of GFP

The sequence of *m-gfp* was mutated by PCR in the presence of limiting nucleotide concentrations. The template plasmid was pBSm-gfp4, a derivative of TU#65 (Chalfie *et al.*, 1994 Science 263, 802-805) in which *gfp* has been replaced with *m-gfp*. The primers used were the T3 and T7 primers (New England Biolabs) that are complementary to the flanking T3 and T7 promoters present in the vector sequence. Four separate reactions (30 cycles of 30 seconds at 94°C, 30 seconds at 55°C and 1 min at 72°C using Taq DNA Polymerase from Promega) were carried out, each with the concentration of a different nucleotide reduced from 200 µM to 20 µM. The amplified fragments were pooled, cleaved with *Kpn*I and *Eco*RI and cloned downstream of the *lac* promoter of pBluescript II KS (+) (Stratagene).

The mutant library thus obtained was transformed into *E. coli* strain XL1-Blue (Stratagene) and incubated overnight at 37°C on TYE agar containing 50 µg/ml ampicillin and 1 mM IPTG. Colonies were illuminated with a long wavelength UV lamp (UVP Model B 100 AP) and visually screened for increased fluorescence. The coding regions of two of the brightest mutant genes (*m-gfpA* and *m-gfpB*) thus identified, as well as that of *m-gfp*, were amplified by PCR (30 cycles of 1 min at 94°C, 1 min at 55°C and 1 min at 72°C using VENT DNA Polymerase from New England Biolabs) using primers

that generate a *Bam*HI-*Sac*I fragment containing the *gfp* coding sequence downstream of a phage Shine-Dalgarno sequence and a plant translation initiation context sequence. The forward primer (5'*Bam*-GFP) was

5'-GGCGGGATCCAAGGAGATAACAATGAGTAAAGGAGAAGAACTTTCACT-3'

(Seq ID No. 5, *Bam*HI site underlined, Shine-Dalgarno sequence in italics and translation initiation context sequence in bold) and the reverse primer (GFP-3'*Sac*) was 5'-GGCGAGCTTATTGTATAGTCATCCATGCC-3' (Seq ID No. 6, *Sac*I site underlined and GFP stop codon in bold). The amplified fragments obtained from the three reactions were cleaved with *Bam*HI and *Sac*I and cloned downstream of the *lac* promoter of pUC119 (Vieira and Messing, 1987 Methods Enzymol. 153, 3-11).

The positions of the mutations responsible for the bright phenotypes of *m-gfpA* and *m-gfpB* were then localised by recombination of the mutant genes with *m-gfp*. The pUC119 derivatives containing *m-gfpA* and *m-gfpB* were cleaved with either *Bam*HI and *Nco*I, *Nco*I and *Cla*I, or *Cla*I and *Sac*I. The restriction fragments were gel purified and ligated to the *m-gfp* pUC119 derivative that had been cleaved with the same combination of enzymes and gel purified. These and the parent constructs were introduced into XL1-Blue cells and incubated overnight at 37°C on agar plates containing ampicillin and IPTG. Comparison of the fluorescence of colonies containing the various constructs revealed that the mutation(s) responsible for the bright phenotypes of both *m-gfpA* and *m-gfpB* were contained within the 336 bp *Cla*I-*Sac*I fragment at the 3' end of the gene. These fragments were cloned into the phage vector M13mp19 (New England Biolabs) and sequenced from the Universal primer using the Sequenase Version 2.0 DNA Sequencing Kit (United States Biochemical Corporation).

Sequencing of the *Cla*I-*Sac*I fragment of *m-gfpB* revealed the presence of a single coding alteration, V163A. This same change was found in combination with a second coding alteration, S175G, in the *Cla*I-*Sac*I fragment of *m-gfpA*. The sequences of mGFP, mGFPB and mGFPA are compared in Figure 6, which shows the location of the two mutations. The S175G change would appear to contribute to the phenotype of GFPA as cells expressing the GFPA protein were clearly more fluorescent than those expressing GFPB (data not shown). Thus, only GFPA was analysed further.

The sequence of the *m-gfp* gene was modified so as to code for the V163A and S175G substitutions described above and the I167T substitution described in the prior art (1994 Heim et al., Proc. Natl. Acad. Sci. USA 91, 12,501-12,504, which substitution inverts the ratio of the 400-475nm excitation peaks), as well as to further alter the codon usage of the gene in order to eliminate potential plant intron sequences generated by the introduction of these mutations. The sequence differences between this modified gene termed, *m-gfp5*, and the original *gfp* gene, and their respective polypeptides are summarised in Fig. 15 (Seq. ID Nos. 5-8).

The *m-gfp5* gene was constructed by PCR amplification (30 cycles of 30 secs at 94°C, 30 secs at 55°C and 30 secs at 72°C using VENT DNA Polymerase) of *m-gfp* using mutagenic primers. The forward primer was an oligo corresponding to nucleotides 445-560 of the *m-gfp5* coding sequence shown in Figure 15 and the reverse primer was GFP-3'Sac. The amplified fragment was cleaved and exchanged with the *AccI-SacI* fragment of *m-gfp* to create *m-gfp5*.

For bacterial expression studies, *BamHI-SacI* PCR fragments containing the *m-gfp*, *m-gfpA* and *m-gfp5* genes were cloned downstream of the *tac* promoter of the expression vector pSE380 (Invitrogen), to give the plasmids pSE-GFP, pSE-GFPA and pSE-GFP5, respectively. Expression from the *tac* promoter of pSE380 is tightly regulated due to the presence on the plasmid of the *lacIq* gene. For yeast expression, the same PCR fragments containing the *m-gfp*, *m-gfpA* and *m-gfp5* genes were inserted downstream of the constitutive *ADH1* promoter of pVT103-U (Vernet et al., 1987), a yeast multicopy episomal plasmid containing the *URA3* selectable marker. The resulting plasmids were pVT-GFP, pVT-GFPA and pVT-GFP5, respectively.

To assess the difference in fluorescence between strains expressing modified GFP and GFPA quantitatively, the inventors introduced expression plasmids pSE-GFP and pSE-GFPA into *E. coli* strain XL1-Blue and measured the fluorescence ($\lambda_{ex} = 397$ nm, $\lambda_{em} = 508$ nm) of equal optical densities of cells at various times following IPTG-induction of protein synthesis at 37°C (Fig. 7). 4.5 hrs after induction, cells expressing m-GFPA were observed to fluoresce approximately 20-fold more intensively than those expressing

m-GFP, a figure which increased to approximately 35-fold by the time the cells had entered stationary phase (9 hrs after induction).

To determine whether the enhanced fluorescence of cells expressing m-GFPA might be due to increased levels of protein expression, total protein was prepared from cells from the 4.5 hr time point and the amount of intracellular GFP estimated by Western blot analysis. As can be seen in Fig. 8, m-GFPA accumulates to a significantly higher level than modified GFP. The "vector" track is a negative control. The numbers on the right of the blot represent the molecular weights of known standards. However, the difference in protein levels as estimated by quantification of band intensities, 2.4-fold, is not nearly enough to account for the approximately 20-fold difference in fluorescence levels observed at this time point. This result suggests that a large proportion of GFP that is expressed in cells at 37°C is non-fluorescent and that the substitutions present in m-GFPA enhance the maturation of the protein to the fluorescent form. Comparison of the growth curves of strains expressing m-GFP and m-GFPA with the growth curve of a non-expressing strain (data not shown) indicated that expression of these proteins does not have any adverse effects on the growth of bacterial cells.

The amino acid substitutions present in m-GFPA suppress the temperature-sensitivity of GFP maturation

Lim and co-workers (Lim *et al.*, cited above) have recently reported that maturation of GFP to the fluorescent form is sensitive to temperature during expression in the yeast *Saccharomyces cerevisiae*. To test whether the same may be true during expression in *E. coli* and whether the substitutions present in m-GFPA enhance maturation by suppressing any such sensitivity, the inventors examined expression of m-GFP and m-GFPA over a range of different temperatures. Strains containing pSE-GFP and pSE-GFPA were induced overnight at temperatures ranging between 25°C and 42°C. For each culture, the fluorescence of equal optical densities of cells was measured and the amount of intracellular GFP determined by Western blot analysis (Table 1). Fluorescence values were then normalised against the amount of GFP present inside cells so as to give a relative measure of the proportion of intracellular GFP that is fluorescent for each culture. The results (Fig. 9) clearly show that the proportion of modified GFP

that is fluorescent steadily decreases with increasing incubation temperature (open columns), indicating that maturation of the protein to the fluorescent form is temperature sensitive. In contrast, the substitutions present in m-GFPA (shaded columns) suppress this sensitivity to temperature, with maturation being optimal at 37°C.

Table 1

Temperature (°C)	Fluorescence (arbitrary units)		Relative amount of intracellular GFP (units)	
	m-GFP	m-GFPA	m-GFP	m-GFPA
25	328.4	722.2	0.29	0.58
30	100.5	541.1	0.21	0.82
37	67.9	2273.0	0.23	1.00 (7.8×10^5)
42	9.2	369.4	0.17	0.44

Investigation into the thermosensitivity of GFP maturation

The post-translational maturation of GFP to the fluorescent form involves a number of steps. The first step, presumably, is folding of the apoprotein into a catalytic conformation that facilitates the novel reactions involved in formation of the chromophore. These reactions consist of cyclisation and oxidation of the tripeptide Ser65-Tyr66-Gly67 to give a p-hydroxybenzylidene-imidazolidinone structure. Once the chromophore has been formed, it is then only fluorescent once GFP has adopted a fold which protects it from solvent effects. In principle, any of these processes could be sensitive to temperature and thus be responsible for the observed thermosensitivity of GFP maturation.

Since the oxidation reaction involved in chromophore formation appears to require molecular oxygen, Heim and co-workers (Heim *et al.*, 1995 *Nature* 373, 663-664) have been able to measure the reaction rate by expressing GFP in *E. coli* under anaerobic conditions and then monitoring the development of fluorescence after admission of air.

To determine whether this reaction might be temperature sensitive and whether the substitutions present in m-GFPA act by enhancing its rate at higher temperatures, we measured the rates of oxidation of m-GFP and m-GFPA at both 25°C and 37°C. For these experiments a yeast expression system was used, which provided for better growth and expression levels than *E. coli* under anaerobic conditions. Strains of *Saccharomyces cerevisiae* MGLD-4a (*MATα, leu2, ura3, his3, trp1, lys2*) containing either pVT-GFP or pVT-GFPA were incubated anaerobically (Becton-Dickinson BBL GasPak Pouch) overnight at 30°C in synthetic drop-out media lacking uracil (Rose *et al.*, 1990 "Methods in Yeast Genetics. A Laboratory Course Manual", Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA). Following admission of air to the pouch, 1.0 ml of each culture was immediately centrifuged for 1 min at 13,000 rpm and resuspended in 0.5 ml aerated and prewarmed PBS (pH 7.4) containing 8 mM NaN₃ as a metabolic inhibitor. Cell suspensions were placed immediately into pre-warmed cuvettes held within the fluorimeter carousel and the time course of fluorescence ($\lambda_{ex} = 397$ nm, $\lambda_{em} = 508$ nm) development measured.

As reported previously by Heim *et al.*, each oxidation proceeded as a simple first order reaction (Fig. 10). Figure 10 shows the rate of fluorescence development for cultures expressing modified GFP at 25°C (crosses) or 37°C (triangles), or cultures expressing GFPA at 25°C (squares) or 37°C (circles).

The time constant measured for the oxidation of m-GFP at 37°C (5.9 ± 0.1 min) was found to be approximately 3-fold faster than that measured at 25°C (16.2 ± 0.3 min), indicating that the post-translational oxidation of the GFP chromophore is not the step responsible for the temperature sensitivity of maturation. In confirmation of this conclusion, the time constants derived for m-GFPA at both 25°C and 37°C (22.5 ± 1.4 min and 18.1 ± 0.4 min, respectively) were actually slower than those measured for m-GFP.

Heim and co-workers have also reported that some GFP forms non-fluorescent inclusion bodies during expression in *E. coli*, indicating that not all GFP folds properly under these conditions. To determine whether the proper folding of m-GFP might be temperature

sensitive and whether the substitutions present in m-GFPA act by enhancing proper folding at increased temperatures. the inventors examined the solubilities of the two proteins during expression in *E. coli* at 25°C and 37°C. Bacterial cells expressing m-GFP or m-GFPA were grown overnight at either 25°C or 37°C. lysed, and the soluble and insoluble fractions separated by centrifugation.

Specifically, cells containing pSE-GFP or pSE-GFPA were grown in 1.5 ml of 2xTY broth to an absorbance of 0.2 at 600 nm and then induced overnight with 0.2 mM IPTG. The cultures were centrifuged at 13,000 rpm for 2 min, resuspended in 500 µl 50 mM Tris-HCl (pH 8.0), 2 mM EDTA, 100 µg/ml lysozyme, 0.1% Triton X-100 and incubated at 30°C for 15 min. Cells were then lysed by sonication (5 x 15 secs) using a Heat Systems (Model CL4) sonicator and centrifuged at 13,000 rpm for 15 min at 4°C. The supernatant (soluble fraction) was removed and stored at -70°C until used. The pellet (insoluble fraction) was washed once with 500 µl 50 mM Tris-HCl (pH 8.0), 10 mM EDTA, 0.5% Triton X-100, resolubilised for 1 hr at room temperature in 500 µl resolubilisation buffer (8 M urea, 0.1 M NaH₂PO₄, 10mM Tris-HCl, pH 6.3) and stored at -70°C until used. The amount of m-GFP or m-GFPA present in each fraction was then estimated by Western blot analysis (Fig. 11).

SDS-PAGE and Western blot analysis were carried out according to normal procedure (Sambrook *et al.*, 1989 "Molecular Cloning. A Laboratory Manual". Cold Spring Harbor Laboratory Press, Cold Spring Harbor, USA). Primary antibodies were polyclonal rabbit anti-GFP (generous gift of S. Santa-Cruz) used at a dilution of 1/2,000. Antibodies were detected with iodinated Protein A (Amersham) and bands visualised and quantified using a Molecular Dynamics Phosphorimager. In Figure 11, insoluble and soluble fractions are denoted by the letters I and S respectively, with cultures grown at 25°C shown on the left, and cultures grown at 37°C shown on the right. In all cases, fluorescence was found almost exclusively in the soluble fraction. At 25°C, both m-GFP and m-GFPA were found predominantly in the soluble fraction, indicating that proper folding of both proteins is efficient at this temperature. At 37°C, however, the majority of m-GFP was found as aggregated protein in the insoluble fraction, whereas most of m-GFPA was still present in the soluble fraction. This result indicates that the temperature

sensitivity of m-GFP maturation is due primarily to improper protein folding at higher temperatures and that this defect is suppressed by the amino acid substitutions present in m-GFPA.

To gain information on which species in the maturation pathway of GFP aggregates at higher temperatures, the inventors made use of the characteristic absorption of the GFP chromophore in either the mature (Ward & Bokman, 1982 Biochemistry 21, 4535-4540) or chemically reduced state (Inouye & Tsuji 1994 FEBS Lett. 351, 211-214). If the aggregating species has already undergone the cyclisation reaction, GFP isolated from inclusion bodies should show this characteristic absorption. To facilitate the purification of protein for absorbance measurements, the inventors fused a polyhistidine tag to the C-terminus of m-GFP.

Histidine-tagging was achieved by the addition of 6 histidine codons to the 3' ends of the modified *gfp* genes by PCR. The genes were amplified using 5'Bam-GFP as the forward primer and the oligo

5'- GCCGAGCTCTTAGTGGTGGTGGTGGTGGT
 TTTGTATAGTTCATCCATGCC -3'

(Seq ID No. 7, *SacI* site underlined, histidine codons in bold) as the reverse primer. The amplified fragments were cleaved with *BamHI* and *SacI* and cloned downstream of the *tac* promoter of pSE380 (Invitrogen) to give the expression plasmids pSE-GFPHis, pSE-GFPAHis and pSE-GFP5His respectively.

For the purification of histidine-tagged GFP for absorbance measurements, soluble and insoluble fractions of cells containing pSE-GFPHis grown at 25°C and 37°C, respectively, were prepared as described previously. GFP was purified from the fractions on Ni-chelate columns using the Ni-NTA Spin Kit (Qiagen). Purification from the soluble fraction was carried out according to the protocol for the purification of histidine-tagged proteins under native conditions. After clearance of cellular debris from the insoluble fraction by centrifugation at 13,000 rpm for 30 min. purification was carried out according to the protocol for purification of histidine-tagged proteins under denaturing conditions. except that the protein was eluted with resolubilisation buffer containing 250

mM imidazole.

For the purification of histidine-tagged proteins for fluorescence spectroscopy, cells were grown in 100 ml of 2xTY broth at 37°C to an absorbance of 0.2 at 600 nm and then induced overnight with 0.5 mM IPTG. Cells were harvested by centrifugation at 6,000 rpm for 10 min and lysed by resuspension in 4 ml 20 mM Tris-HCl (pH 7.9), 500 mM NaCl, 5 mM imidazole, 0.1% sarkosyl, 0.1% deoxycholate, 2.25 M Guanidine-HCl. Nucleic acids were precipitated by the addition of 5ml isopropanol and removed by centrifugation at 10,000 rpm for 10 min. Fluorescent histidine-tagged proteins were purified from the supernatant on Ni-chelate columns (Qiagen) and eluted with 2ml of 20mM Tris-HCl (pH 7.9), 500 mM NaCl, 150 mM imidazole. For all purifications, protein purity was assayed by SDS-PAGE and found to be >95%. Protein concentrations were determined by Bradford assay (Bio-Rad Protein Assay kit) using bovine serum albumin as a standard.

Absorbance spectra were recorded on a Cary 3 UV-Visible Spectrophotometer (Varian) at 25°C. The optical pathlength was 1 cm. Fluorescence spectra were recorded on a Hitachi F-4500 fluorimeter at 25°C using 4mm/10mm cuvettes. The bandpass for both the excitation and the emission monochromators was 5 nm, the scan speed 240 nm per min and the response time automatically adapted by the device. All spectra were corrected following the supplier's procedure for calibration of the fluorimeter using Rhodamine-B as standard. Emission spectra were recorded at a fixed wavelength of the excitation maximum, excitation spectra at a fixed wavelength of the emission maximum.

Histidine-tagging of GFP did not detectably affect the temperature sensitivity of maturation of the protein (data not shown). The absorption spectra of equal concentrations of denatured protein from the two preparations were then recorded, as described above. The results are shown in Figure 12, which is a graph of absorbance against wavelength (nm).

As can be seen in Fig. 12, denatured fluorescent protein derived from the soluble fraction of cells grown at 25°C shows a characteristic absorption peak similar to that of the m-

GFP chromophore at acid pH (Ward & Bokman, cited above). On the other hand, protein purified from inclusion bodies of cells grown at 37°C shows no such absorption, indicating that the aggregating species has not formed a chromophore. Taken together, the results presented above indicate that the temperature sensitivity of m-GFP maturation is due primarily to the failure of the unmodified apoprotein to fold into its catalytically active conformation at higher temperatures. Furthermore, the amino acid substitutions present in m-GFPA suppress this defect by enhancing proper folding at elevated temperatures.

If thermosensitivity of m-GFP maturation observed in the yeast *Saccharomyces cerevisiae* is also a result of the thermosensitivity of apoprotein folding, it should be suppressed by the substitutions present in m-GFPA. To test this prediction, the inventors incubated strains of *cerevisiae* containing either pVT-GFP or pVT-GFPA on agar plates at either 25°C or 37°C. As can be seen in Fig. 13, the substitutions present in m-GFPA also suppress the thermosensitivity of m-GFP expression in yeast. This result indicates that the temperature-dependent mis-folding of the m-GFP apoprotein is not simply an artefact of an *E. coli* overexpression system, but is also the basis for the thermosensitivity of m-GFP maturation in a heterologous eukaryotic system.

Modification of the fluorescence spectra of m-GFPA

Fluorescence spectroscopy of purified histidine-tagged m-GFP and m-GFPA revealed that the fluorescence spectra of m-GFPA are essentially unchanged from those of m-GFP except for a decrease in the amplitude of the 475 nm excitation peak relative to the amplitude of the 400 nm excitation peak (Fig. 14). Although this spectral change is advantageous for applications which utilise 400 nm excitation, it is also detrimental for those which utilise 475 nm excitation. For many experiments, the ideal spectral variant would be a protein which could be efficiently excited at either of these wavelengths. This characteristic would afford greater flexibility with regard to the range of applications in which the protein could be used.

Recently, it has been demonstrated that, as observed here, the relative amplitudes of the excitation peaks of GFP can be altered by means of mutagenesis (Ehrig *et al.*, 1995

FEBS Lett. 367, 163-166; Delagrange *et al.*, 1995 Bio/Technology 13, 151-154). A number of these mutations, like the substitutions present in m-GFPA, are located in the C-terminal region of the protein. It has been hypothesised that, in the three-dimensional structure of GFP, the C-terminal region is close to the chromophore and that mutations in this region can affect the microenvironment of the chromophore so as to influence the equilibrium between the two tautomeric forms of the chromophore that are responsible for the two excitation peaks (Heim *et al.*, 1994 & Ehrlig *et al.*, 1995). One of these mutations, I167T, inverts the ratio of the 400 nm to 475 nm excitation peak heights. If the effects of mutations in the C-terminal region of GFP on the spectroscopic state of the chromophore are additive, then it is possible that combination of the I167T substitution with the substitutions present in GFPA might increase the amplitude of the 475 nm peak relative to the 400 nm peak.

Histidine-tagged m-GFP5 was purified and its excitation and emission spectra analysed by fluorescence spectroscopy. As can be seen in Fig. 14, m-GFP5 has two excitation peaks (maxima at 395 nm and 473 nm) of almost exactly equal amplitude and an emission spectrum largely unchanged from that of m-GFP. To determine whether m-GFP5 has retained the thermotolerant phenotype of GFPA, bacterial cells containing pSE-GFP or an expression plasmid containing *m-gfp5* (pSE-GFP5) were induced with IPTG for 5 hours at 37°C. The fluorescence ($\lambda_{ex} = 395$ nm or 473 nm, $\lambda_{em} = 507$ nm) of equal optical densities of cells was then measured. Cells expressing m-GFP5 were observed to fluoresce 39-fold more intensely than cells expressing m-GFP when excited at 395 nm and 111-fold more intensely when excited at 473 nm. These results indicate that m-GFP5 has not only retained the thermotolerant phenotype of m-GFPA, but has improved upon it.

Further modification of *mgfp5* was achieved. Two synthetic oligonucleotides were made, to act as mutagenic PCR primers to add an in-frame *Eco*RI site at the 5' end of the gene and to add a sequence coding for the amino acid tag HDEL at the C terminal of the protein (which tag acts as an endoplasmic reticulum localisation signal).

The PCR mutagenised sequence was then used in a three-way ligation reaction with

*Bam*HI/*Sac*I - cut vector and a pair of synthetic oligonucleotides with *Bam*HI/*Eco*RI ends. The synthetic oligos had the sequences:

5' GGC GGA TCC AAG GAG ATA TAA CAA TGA AGA CTA
ATC TTT TTC TCT TTC TCA TCT TTT CAC 3' (Seq ID No. 8) and

5' GCC GAA TTC GGC CGA GGA TAA TGA TAG
GAG AAG TGA AAA GAT GAG AAA GA 3' (Seq ID No. 9)

The oligos were annealed, extended with Klenow polymerase and cut with *Bam*HI/*Eco*RI. When ligated with the *m-gfp5HDEL* gene, they introduced the signal sequence from *Arabidopsis* chitinase at the 5' end of the coding sequence. The nucleotide sequence of the resulting modified gene (*m-gfp5-ER*), and the amino acid sequence of its polypeptide product (Seq ID No.s 10 and 11 respectively), are shown in Figure 16. The nucleotides encoding the signal sequence are shown in upper case letters, whilst the rest of the sequence is in lower case letters. The C terminal HDEL tag on the protein is apparent.

The modified gene, when expressed in *Arabidopsis*, gave highly efficient concentration of GFP-mediated fluorescence in the endoplasmic reticulum (Figure 17). Referring to Figure 17, the panels illustrate confocal micrographs of 5-day old *A. thaliana* seedlings expressing m-GFP (panels A-D) or m-GFP5-ER (panels E-H), imaged at 395nm excitation wavelength. Panels A & E are sections of the junction between the hypocotyl and the cotyledon (bar = 25μM); panels B & F show hypocotyl epidermal cells (bar = 10μM); panels C & G show median longitudinal sections of root tips (bar = 25μM); and panels D & H show roots, with nucleoplasmic accumulation of mGFP (in D) and retention in the endoplasmic reticulum of mGFP5-ER (in H), (bar = 10μM).

DISCUSSION

GFP expression in plants

The objective of this work was to begin the development of *gfp* for use as a genetic marker in transformation and as a reporter for localised gene expression in *Arabidopsis*

and other plants. In order to successfully employ the *gfp* cDNA in plants, three major steps need to be addressed.

- (1) The GFP apoprotein must be produced in suitable amounts within the plant cells.
- (2) The apoprotein must undergo efficient post-translational cyclisation and oxidation to produce the mature GFP.
- (3) The fluorescent protein may need to be suitably targeted within the cell, to allow efficient post-translational processing, safe accumulation to high levels, or to allow easier distinction of expressing cells.

The inventors have shown that expression of the jellyfish *gfp* cDNA in *Arabidopsis* is curtailed by aberrant splicing, with an 84 nucleotide intron being efficiently excised from within the GFP coding sequence. The recognition of introns in plant pre-mRNAs primarily requires conserved sequences found adjacent to the 5' and 3' splice sites, which are related to those found in other eukaryotes, and, atypically, a high A:U content within the intron. The inventors altered potential recognition sequences at the 5' splice site, and decreased the A:U content of the cryptic intron by *in vitro* mutagenesis to produce a modified *m-gfp* gene which was successfully expressed in transgenic *Arabidopsis* plants. It is likely that this *m-gfp* gene will be useful for expression studies in other plants, which appear to share similar features involved in intron recognition.

It is also possible that aberrant splicing may interfere with GFP expression in other organisms. However, introns found in yeast possess a requirement for conserved sequences located at the branch point, and introns found in animal cells (including jellyfish) share a conserved polypyrimidine tract adjacent to the 3' splice site. The lack of these additional features may allow correct processing of the *gfp* mRNA in fungal and animal cells.

With expression of the *m-gfp* gene in *Arabidopsis*, it has now been shown that the apoGFP readily undergoes maturation, and that the fluorescent form of the protein accumulates in transformed cells. Transformed cells were often intensely fluorescent, and were easily detectable by eye using a long-wave UV lamp. However it proved difficult to efficiently regenerate fertile plants from the brightest transformants, with cells

remaining as a highly fluorescent callus or mass of shoots after several months of culture. It is possible that high levels of GFP expression were mildly toxic or interfered with differentiation, due perhaps to the fluorescent or autocatalytic properties of the protein. In the natural situation, in jellyfish photocytcs where high levels of GFP are well tolerated, the protein is found sequestered in microbody-like lumisomes. In contrast, the mature protein is found throughout the cytoplasm and nucleoplasm in transformed *Arabidopsis*. If GFP is a source of fluorescence-related free radicals, for example, it might be advisable to target the protein to a more localised compartment within the plant cell. Appropriate localisation signals are known to those skilled in the art and it should prove possible to incorporate these into the GFP polypeptide without unduly disrupting the fluorescence characteristics of the protein.

The inventors have adapted the green fluorescent protein (GFP) of *Aequoria victoria* for use as a genetic marker in *Arabidopsis thaliana*. Transcripts of the jellyfish GFP coding sequence are mis-spliced in *Arabidopsis*, with an 84 nucleotide intron being efficiently excised. A modified version of the *gfp* sequence has been constructed to destroy this cryptic intron, and to restore proper expression of the protein in plant cells. GFP is mainly localised within the nucleoplasm and cytoplasm within transformed *Arabidopsis* cells, and its presence allows optical sectioning of intact plants using confocal laser scanning microscopy. The modified *gfp* sequence may be useful for directly monitoring gene expression and protein localization at high resolution, and as a simply scored genetic marker in living plants.

A major use for *m-gfp* would be as a replacement for the β -glucuronidase gene, used as a reporter for promoter and gene fusions in transformed plants. Histochemical staining is used to identify cells expressing the GUS gene product, but a fluorescent product can be imaged directly and rapidly. Gene expression and protein localization can be observed in physiologically active cells without a prolonged and lethal staining procedure, and fluorescence microscopy techniques allow the high resolution imaging of GFP-expressing cells. In addition, it becomes feasible to follow dynamic events in living cells and tissues.

High levels of fluorescence intensity are obtained in GFP-transformed bacterial and yeast colonies allowing simple screening for GFP expression with the use of a hand-held UV lamp. Such an assay for gene expression in living plants would be a very useful tool for plant transformation experiments. Many transformation techniques give rise to regenerating tissues which are variable or chimeric, and require testing of the progeny of the primary transformants. Potentially, *m-gfp*-transformed tissues could be monitored using *in vivo* fluorescence, avoiding any need for destructive testing, and the appropriate transformants could be rescued and directly grown to seed. Similarly, *in vivo* fluorescence would be an easily scored marker for field testing in plant breeding, allowing *m-gfp*-linked transgenes to be easily followed.

Use of a confocal laser scanning microscope will allow the clear analysis of plant tissue whole-mounts despite the refractile nature and light scattering (and for some cells, autofluorescent) properties of plant cell walls.

In this study, the inventors have also shown that maturation of GFP in *E. coli* is sensitive to temperature, due primarily to the mis-folding of the apoprotein into inclusion bodies at elevated temperatures. They have also described two mutants, m-GFPA and m-GFP5, whose folding is thermotolerant. Presumably, the characteristic of the GFP apoprotein that causes it to aggregate at higher temperatures and the mechanism by which the mutations present in m-GFPA and m-GFP5 suppress this effect is unknown. However, studies on the effects of mutations on the tendency of some other proteins to aggregate (Thomas *et al.*, 1995 Trends Biochem. Sci. 20, 456-459; Mitraki *et al.*, 1991 Science 253, 54-58; Wetzel 1994 Trends Biotech. 12, 193-198; and Chrunyk *et al.*, 1993 J. Biol. Chem. 268, 18053-18061) suggest a number of possibilities.

The simplest explanation is that the native apoprotein or one of its folding intermediates is thermodynamically unstable and the protein aggregates when in the unfolded state. The substitutions present in the thermotolerant mutants could suppress the characteristic either by increasing the thermodynamic stability of the unstable species or by decreasing its steady state level by increasing the rate of chromophore cyclisation. Alternatively, higher temperatures can allow proteins to overcome the thermodynamic barriers to the

formation of off-pathway folding intermediates which may become kinetically trapped by aggregation. It is possible, therefore, that the substitutions present in the heat-tolerant mutants act by suppressing such a phenomenon by directing folding along the correct pathway at elevated temperatures. It is also possible that the kinetic half life of an aggregation-prone intermediate in the normal apoprotein folding pathway is significantly increased at higher temperatures. Mutations could suppress this characteristic either by decreasing the half life of such an intermediate or by reducing its tendency to aggregate. Finally, the suppressor mutations may favour proper folding at higher temperatures by increasing the affinity of the apoprotein for a molecular chaperone.

To differentiate between these different possibilities, biophysical analyses of variants of GFP and the heat tolerant mutants that cannot undergo cyclisation and are thus trapped as apoproteins will be required. However, the observation that the substitutions present in m-GFPA increase the T_m of mature GFP by 4.0°C (data not shown) provides preliminary evidence that they may act by increasing the thermodynamic stability of the native apoprotein. As well as enhancing proper folding, the substitutions present in m-GFPA contribute to the bright phenotype of the mutant protein by facilitating its accumulation to higher levels than m-GFP (Table 1). This observation indicates that GFP is turned over more rapidly than m-GFPA, probably because partially or mis-folded m-GFP apoprotein that does not aggregate would be degraded by the cellular proteolytic machinery.

The inventors have shown that oxidation of the GFP chromophore does not contribute to the temperature sensitivity of maturation by measuring the reaction rate in yeast cells at both 25°C and 37°C (Fig. 3). An interesting point arising from this experiment is that the time constants derived for m-GFP at both 25°C and 37°C (5.9 ± 0.1 min and 16.2 ± 0.3 min, respectively) are significantly faster than the 120 min estimated for the oxidation of GFP in bacteria by Heim *et al.* This observation may reflect a difference in the physiological states of yeast and bacterial cells following anaerobic growth or perhaps the presence of a catalysing factor in yeast cells.

Nevertheless, these results suggest that the oxidation of the GFP chromophore has the

capacity to proceed at a much higher rate than previously thought. Therefore, in some cases, the factor which limits how quickly fluorescent can be observed following protein synthesis may be the efficiency with which the apoprotein folds rather than the time taken for oxidation of the chromophore.

Examination of the fluorescence spectra of the m-GFPA revealed a decrease in the amplitude of the 475nm excitation peak relative to the amplitude of the 400nm excitation peak. This result indicates that mutations in the C-terminal region of GFP are able to modulate the spectroscopic state of the chromophore by affecting its local environment within the protein. The inventors have utilised this phenomenon to engineer the fluorescence spectra of m-GFP5 by introducing a third substitution, I167T, into the C-terminal region of m-GFPA. m-GFP5 has two excitation peaks (maxima at 395nm and 473nm) of almost exactly equal amplitude and is thus ideal as a multi-purpose spectral variant which can be used for applications requiring both UV and blue excitation. Since the substitutions present in m-GFPA and m-GFP5 affect the environment of the chromophore, it is likely that they also influence the intrinsic brightness of the mutant proteins by affecting the extinction coefficients and/or quantum yields of the chromophore at the two excitation wavelengths. However, to measure an extinction coefficient accurately, one must be certain that every GFP molecule in a given sample is mature. The results presented here suggest that, even in a soluble fraction, there may be appreciable amounts of mis-folded or non-fluorescent apoprotein. In support of this notion, the ratio of the absorbance of the chromophore to that of the aromatic amino acids of histidine-tagged m-GFP purified from the soluble fraction of bacterial cells grown at either 25°C (Fig. X+4) or 37°C (Inouye & Tsuji, cited above) is approximately 0.4. Since this value is in excess of 1.0 for either native or acid-denatured GFP isolated directly from the *Aequorea* jellyfish, it would appear that more than half of the recombinant GFP in a soluble fraction does not have a chromophore.

These observations suggest that extinction coefficients and quantum yields may be difficult to measure unambiguously for recombinant forms of GFP. Therefore, great care must be taken when interpreting the effects of mutations that alter the brightness of GFP. For example, a number of mutations in and near the chromophore of GFP have recently

been described that cause significant shifts in the excitation and/or emission spectra of the protein. A subset of these mutations that alter the tyrosine residue at position 66 to tryptophan, histidine or phenylalanine progressively blue-shift the excitation and emission spectra. However, these mutant proteins are much less fluorescent than GFP, a phenomenon which has been attributed to them having sub-optimal extinction coefficients and/or quantum yields due to the poor fit of the alternative amino acids into the central cavity normally occupied by the tyrosine residue. It is possible, however, that the observed low fluorescence of these mutants is due to detrimental effects of the substitutions on folding and/or chromophore formation, resulting in the presence of large amounts of non-fluorescent protein in soluble fractions. Therefore, it is feasible that the proper maturation of these mutants might be enhanced by the introduction of the amino acid substitutions present in m-GFPA or m-GFP5. Indeed, expression of a protein containing the Y66H mutation in combination with the substitutions present in m-GFPA in *E. coli* at 37°C resulted in a 29-fold increase in fluorescence. The fluorescence spectra of this hybrid protein were also unchanged from those of the Y66H mutant alone (data not shown). Therefore, it is foreseeable that the substitutions present in m-GFPA and m-GFP5 may be combined with these and other pre-existing spectral mutations in the chromophore of GFP to produce a range of spectral variants with greatly improved maturation characteristics.

As well as in *E. coli*, maturation of m-GFP appears to be thermosensitive in the yeast *Saccharomyces cerevisiae* and in mammalian cells. Therefore, it appears that the sensitivity of apoprotein folding to temperature may be a ubiquitous phenomenon. Indeed, it is interesting to note that the brightness of m-GFP in *Arabidopsis thaliana* is markedly increased by its retention in the endoplasmic reticulum (see accompanying paper), where a high concentration of chaperonins may enhance proper folding. It is unlikely, however, that the folding defect of m-GFP would manifest itself in the same way in systems where lower expression levels mean that aggregation may not occur to the same extent as in an *E. coli* overexpression system. Rather, it is more likely that partially or mis-folded apoprotein would, if it did not aggregate heavily, become a target of the cellular proteolytic machinery. Therefore, rather than becoming kinetically trapped by aggregation, improperly folded apoprotein would be depleted by degradation. In

support of this notion, accumulation in yeast of a GFP-nucleoplasmin fusion protein expressed from two gene copies steadily decreases with increasing incubation temperature.

Moreover, if the sensitivity of m-GFP apoprotein folding to temperature is a ubiquitous phenomenon, then use of the thermotolerant mutants described here should result in improved expression in a wide range of experimental systems. Indeed, in this work the inventors have demonstrated that the substitutions present in m-GFPA are capable of suppressing the thermosensitivity of m-GFP expression in the yeast *Saccharomyces cerevisiae*. Expression of m-GFPA has also been observed to give rise to significantly increased fluorescence in *Drosophila melanogaster* embryos incubated at 25°C (A Brand, personal communication) and expression of m-GFP5 fused to endoplasmic reticulum retention signals has been observed to result in high levels of fluorescence in *Arabidopsis thaliana* (data not shown). Most strikingly, expression of both m-GFPA and m-GFP5 has been found to result in greatly increased levels of fluorescence in mammalian cells. Therefore, we anticipate that the thermotolerant mutants described in this work and spectral variants derived from them will be of great benefit for expression in many experimental systems, particularly those such as mammalian cells that utilise higher incubation temperatures.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: Medical Research Council
- (B) STREET: 20 Park Crescent
- (C) CITY: London
- (E) COUNTRY: United Kingdom
- (F) POSTAL CODE (ZIP): W1N 4AL
- (G) TELEPHONE: (0171) 636 5422
- (H) TELEFAX: (0171) 323 1331

(ii) TITLE OF INVENTION: Improvements in or Relating to Gene Expression

(iii) NUMBER OF SEQUENCES: 11

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0. Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 50 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GGCGGGATCCA AGGAGATATA ACAATGAGTA AAGGAGAAGA ACTTTTCACT

50

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GGCGAGCTCT TATTTGTATA GTTCATCCAT GCC

33

(2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 122 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GATCATATGA AGCGGCACGA CTTCTTCAAG AGGCCATGC CTGAGGGATA CGTGCAGGAG	60
AGGACCATCT TCTTCAGGA CGACGGGAAC TACAAGACAC GTGCTGAAGT CAAGTTTGAG	120
GG	122

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 126 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GATGTATAACG TTGTGGAGT TGTAGTTGTA TTCCAATTG TGGCCGAGGA TGTTTCCGTC	60
CTCCTTGAAA TCGATTCCT TAAGCTCGAT CCTGTTGACG AGGGTGTCTC CCTCAAACCTT	120
GACTTC	126

(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 50 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GGCGGATCCA AGGAGATATA ACAATGAGTA AAGGAGAAGA ACTTTCACT	50
---	----

(2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

36

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GGCGAGCTCT TATTTGTATA GTTCATCCAT GCC

33

(2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 51 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GCCGAGCTCT TAGTGGTGGT GGTGGTGGTG TTTGTATAGT TCATCCATGC C

51

(2) INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 60 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GGCGGATCCA AGGAGATATA ACAATGAAGA CTAATCTTT TCTCTTCTC ATCTTTCAC

60

(2) INFORMATION FOR SEQ ID NO: 9:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 50 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GCCGAATTCTG GCCGAGGATA ATGATAGGAG AAGTGAAAAG ATGAGAAAGA

50

(2) INFORMATION FOR SEQ ID NO: 10:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 792 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..789

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

ATG AAG ACT AAT CTT TTT CTC TTT CTC ATC TTT TCA CTT CTC CTA TCA Met Lys Thr Asn Leu Phe Leu Phe Leu Ile Phe Ser Leu Leu Leu Ser 1 5 10 15	48
TTA TCC TCG GCC GAA TTC AGT AAA GGA GAA GAA CTT TTC ACT GGA GTT Leu Ser Ser Ala Glu Phe Ser Lys Gly Glu Glu Leu Phe Thr Gly Val 20 25 30	96
GTC CCA ATT CTT GTT GAA TTA GAT GGT GAT GTT AAT GGG CAC AAA TTT Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe 35 40 45	144
TCT GTC AGT GGA GAG GGT GAA GGT GAT GCA ACA TAC GGA AAA CTT ACC Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr 50 55 60	192
CTT AAA TTT ATT TGC ACT ACT GGA AAA CTA CCT GTT CCA TGG CCA ACA Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr 65 70 75 80	240
CTT GTC ACT ACT TTC TCT TAT GGT GTT CAA TGC TTT TCA AGA TAC CCA Leu Val Thr Thr Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro 85 90 95	288
GAT CAT ATG AAG CGG CAC GAC TTC TTC AAG AGC GCC ATG CCT GAG GGA Asp His Met Lys Arg His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly 100 105 110	336
TAC GTG CAG GAG AGG ACC ATC TTC TTC AAG GAC GAC GGG AAC TAC AAG Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys 115 120 125	384
ACA CGT GCT GAA GTC AAG TTT GAG GGA GAC ACC CTC GTC AAC AGG ATC Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile 130 135 140	432
GAG CTT AAG GGA ATC GAT TTC AAG GAG GAC GGA AAC ATC CTC GGC CAC Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His 145 150 155 160	480
AAG TTG GAA TAC AAC TAC AAC TCC CAC AAC GTA TAC ATC ATG GCC GAC Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp 165 170 175	528
AAG CAG AAG AAC GGC ATC AAA GCC AAC TTC AAG ACC CGC CAC AAC ATC Lys Gln Lys Asn Gly Ile Lys Ala Asn Phe Lys Thr Arg His Asn Ile 180 185 190	576
GAA GAC GGC GGC GTG CAA CTC GCT GAC CAT TAT CAA CAA AAT ACT CCA Glu Asp Gly Gly Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro 195 200 205	624

ATT GGC GAT GGC CCT GTC CTT TTA CCA GAC AAC CAT TAC CTG TCC ACA Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr 210 215 220	672
CAA TCT GCC CTT TCG AAA GAT CCC AAC GAA AAG AGA GAC CAC ATG GTC Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val 225 230 235 240	720
CTT CTT GAG TTT GTA ACA GCT GGG ATT ACA CAT GGC ATG GAT GAA Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu 245 250 255	768
CTA TAC AAA CAC GAC GAA CTC TAA Leu Tyr Lys His Asp Glu Leu 260	792

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 263 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

Met Lys Thr Asn Leu Phe Leu Phe Leu Ile Phe Ser Leu Leu Leu Ser 1 5 10 15
Leu Ser Ser Ala Glu Phe Ser Lys Gly Glu Glu Leu Phe Thr Gly Val 20 25 30
Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe 35 40 45
Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr 50 55 60
Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr 65 70 75 80
Leu Val Thr Thr Phe Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro 85 90 95
Asp His Met Lys Arg His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly 100 105 110
Tyr Val Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys 115 120 125
Thr Arg Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile 130 135 140
Glu Leu Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His 145 150 155 160

39

Lys Leu Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp
165 170 175

Lys Gln Lys Asn Gly Ile Lys Ala Asn Phe Lys Thr Arg His Asn Ile
180 185 190

Glu Asp Gly Gly Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro
195 200 205

Ile Gly Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr
210 215 220

Gln Ser Ala Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val
225 230 235 240

Leu Leu Glu Phe Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu
245 250 255

Leu Tyr Lys His Asp Glu Leu
260

Claims

1. A DNA sequence encoding Green Fluorescent Protein (GFP), the sequence being modified relative to the wild type sequence so as to allow for more efficient expression in a plant cell of a functional GFP polypeptide.
2. A DNA sequence according to claim 1, wherein the modification is such as to reduce the probability of an RNA sequence transcribed therefrom being subject to erroneous splicing in a plant cell.
3. A DNA sequence according to claim 2, comprising a plurality of nucleotide substitutions relative to the wild type sequence, the substitutions serving to reduce the excision from the transcribed RNA of the portion corresponding to nucleotides 400-483 of the DNA sequence.
4. A DNA sequence according to claim 2 or 3, wherein the modification serves to decrease the A/U content of the transcribed RNA.
5. A DNA sequence according to any one of claims 2, 3 or 4, wherein the modification serves to decrease the A/U content of the portion of the transcribed RNA corresponding to nucleotides 400-483 of the DNA sequence.
6. A DNA sequence according to any one of the preceding claims, modified so as to cause an amino acid substitution at residue 163 and/or residue 175 relative to the wild type protein sequence.
7. A DNA sequence according to any one of claims 1-6, further comprising a cellular localisation signal directing the encoded GFP to a particular cellular compartment.
8. A DNA sequence according to claim 7, wherein the encoded GFP is directed to the endoplasmic reticulum (ER).

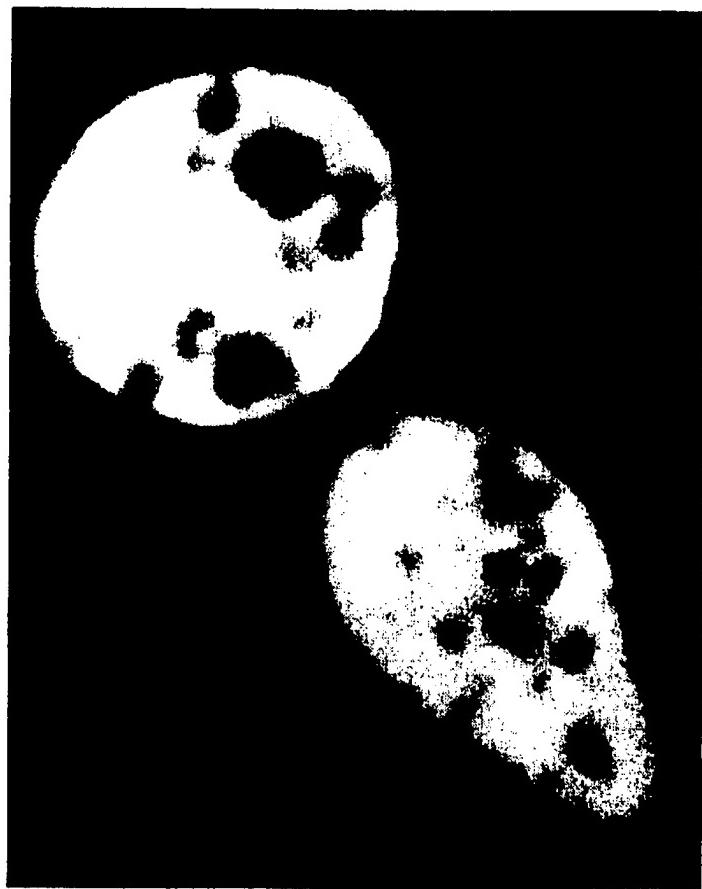
9. A DNA sequence according to claim 8, comprising the *Arabidopsis thaliana* basic chitinase localisation signal.
10. An RNA sequence capable of being transcribed from a DNA sequence according to any one of claims 1-9.
11. A modified GFP polypeptide comprising an amino acid substitution, relative to the wild type protein sequence, at residue 163 and/or 175.
12. A modified GFP according to claim 11, wherein residue 163 is alanine or a related amino acid.
13. A modified GFP according to claim 11 or 12, wherein residue 175 is glycine or a related amino acid.
14. A modified GFP according to any one of claims 11, 12 or 13, comprising one or more further amino acid substitutions relative to the wild type protein sequence.
15. A modified GFP according to any one of claims 11-14, comprising one or more further amino acid substitutions in, or immediately adjacent to, residues 65-67.
16. A modified GFP according to any one of claims 11-15, comprising a localisation signal.
17. A modified GFP according to claim 16, comprising a signal directing the GFP to the endoplasmic reticulum.
18. A modified GFP according to claim 17, comprising the *Arabidopsis thaliana* basic chitinase localisation signal.
19. A nucleic acid sequence encoding a polypeptide according to any one of claims 11-19.

20. A nucleic acid construct comprising a nucleic acid in accordance with any one of claims 1-9 or claim 19.
21. An expression vector according to claim 20.
22. A construct according to claim 20 or 21, wherein the sequence encoding GFP is linked in frame to a sequence of interest.
23. A method of detecting the expression in a host cell of a sequence of interest, comprising causing the sequence of interest to be operably linked in frame with a nucleic acid sequence according to any one of claims 1-9 or claim 19, in a nucleic acid construct according to claim 20, introducing the construct into the host cell, and monitoring for fluorescence characteristic of the modified GFP.
24. A method of screening a plurality of host cells for those which have taken up a nucleic acid sequence of interest, comprising mixing a construct according to claim 20 comprising the sequence of interest with the host cells in suitable conditions, maintaining the host cells under appropriate conditions for a sufficient length of time so as to allow expression of modified GFP, and detecting those cells which exhibit modified GFP-mediated fluorescence.

1 / 21

BamH I	
Sac I	

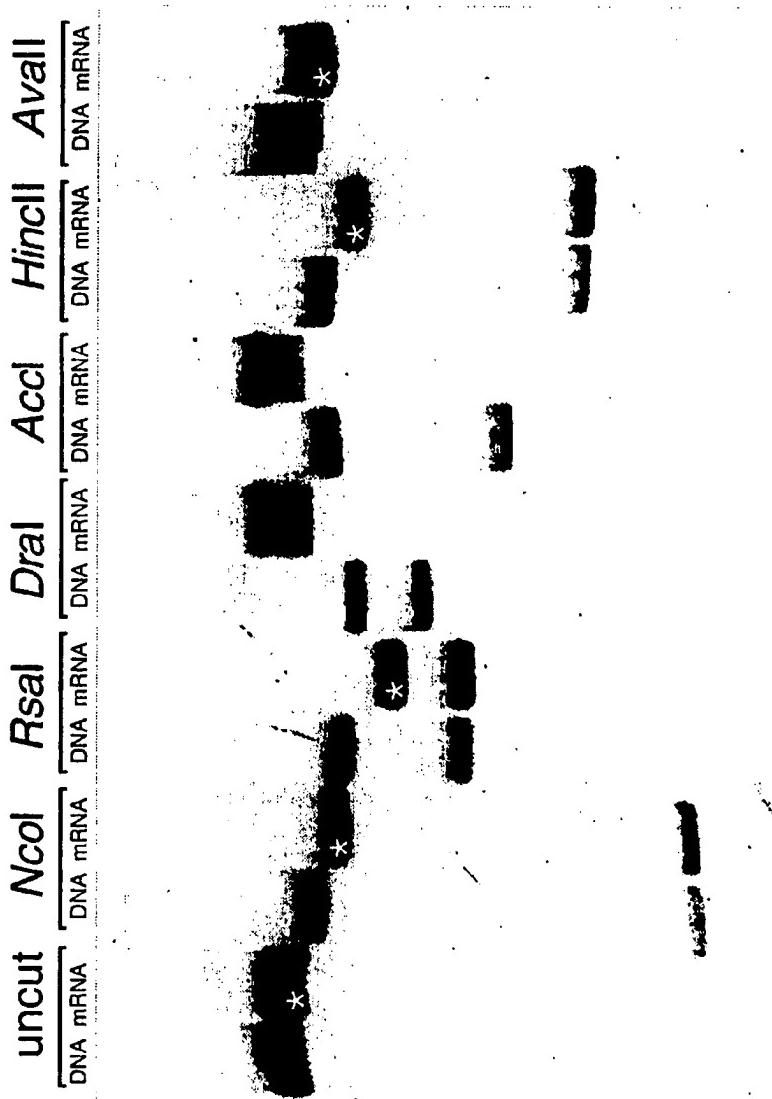
Fig 1A

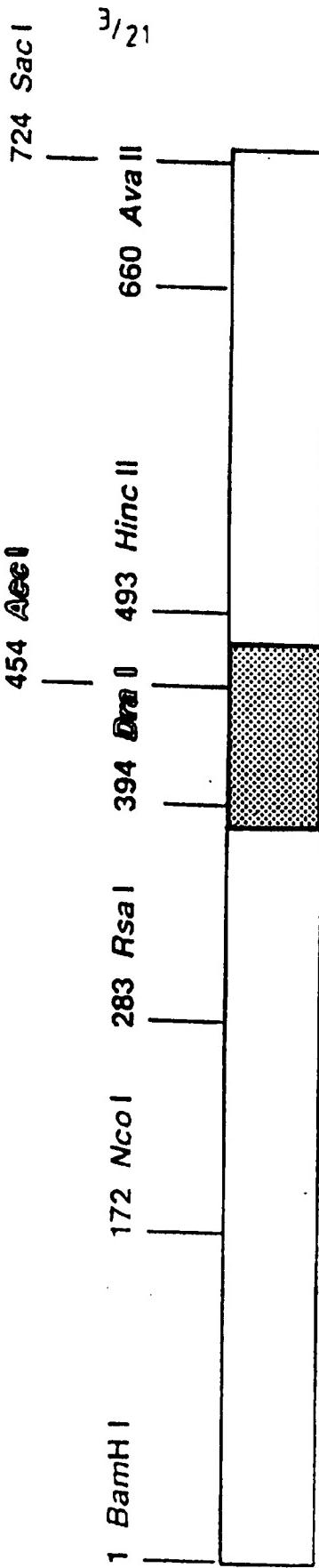


1B
Fig.

2/21

Fig. 2A





GFP coding sequence

SUBSTITUTE SHEET (RULE 26)

Fig. 2B

4/21

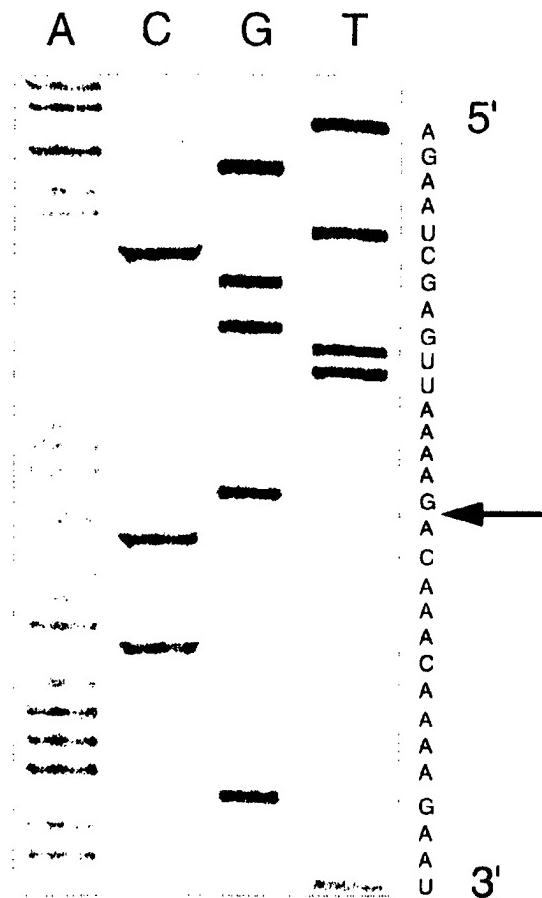


Fig 3A

The diagram illustrates a GFP gene construct. At the top left is a label "5' splice site" with a downward-pointing arrow. In the center is a label "84 base intron". At the top right is a label "3' splice site" with a downward-pointing arrow. Below these labels is a sequence of nucleotides: "...AG GUAUUGA...UCAUGGCAG AC...". Centered below the sequence is the label "GFP sequence".

5' splice site 3' splice site

 ...AG GUAAGU.....AG G....
plant consensus

Fig. 3B

NdeI

201/61	gtc act act ttc tct tat ggt gtt caa tgc ltt tca aga tac cca gat cac atg aaa cgg gtc act act ttc tct tat ggt gtt caa tgc ttt tca aga tac cca gat cac atg aac cgg v T F S Y G C F R S R Y P D H M K R	231/71	
261/81	cat gac ttt ttc aag agt gcc atg ccc gaa ggt tat gta cag gaa aga act ata ttt ttc cac gac ttc ttc aag agc gcc atg cct gag gga tac gtg gag agg acc atc ttc ttc H D F F K S A M P E G Y V Q E R T I F F	291/91	
321/101	aaa gat gac ggg aac tac aag aca cgt gct gaa gtc aag ttt gaa ggt gat <u>acc ctt gtt</u> aaG gac gac ggg aac tac aag aca cgt gct gaa gtc aag ttt gaa gac acc ctc gtc R D G N Y K T R A E V K F E G D T L V	351/111	S/21
381/121	<u>aat aga atc qaq tta aaa qgt att qat ttt aaa gaa gat qga aac att ctt gga cac aaa</u>	411/131	
aac agg atc gag ctt aag gga atc gat tcc aag gag gac gga aac atc ctc ggc cac aag N R I E L K G I D F K E D G N I L G H K			
441/141	<u>ttg qaa tac aac tat aac tca cac aat ata tac atq gca gac aaa caa aag aat gga</u>	AccI 471/151	
L E Y N S H N V I M A D K Q K N G			

top sequence = *Aequorea victoria GFP*, lower sequence = *mGFP4*
 Intron sequences are underlined and the cryptic splice junctions are arrowed.
 Mutated nucleotides are shown outlined. Nucleotide and amino acid
 numbering starts at the initiation codon.

Fig. 4

6/21

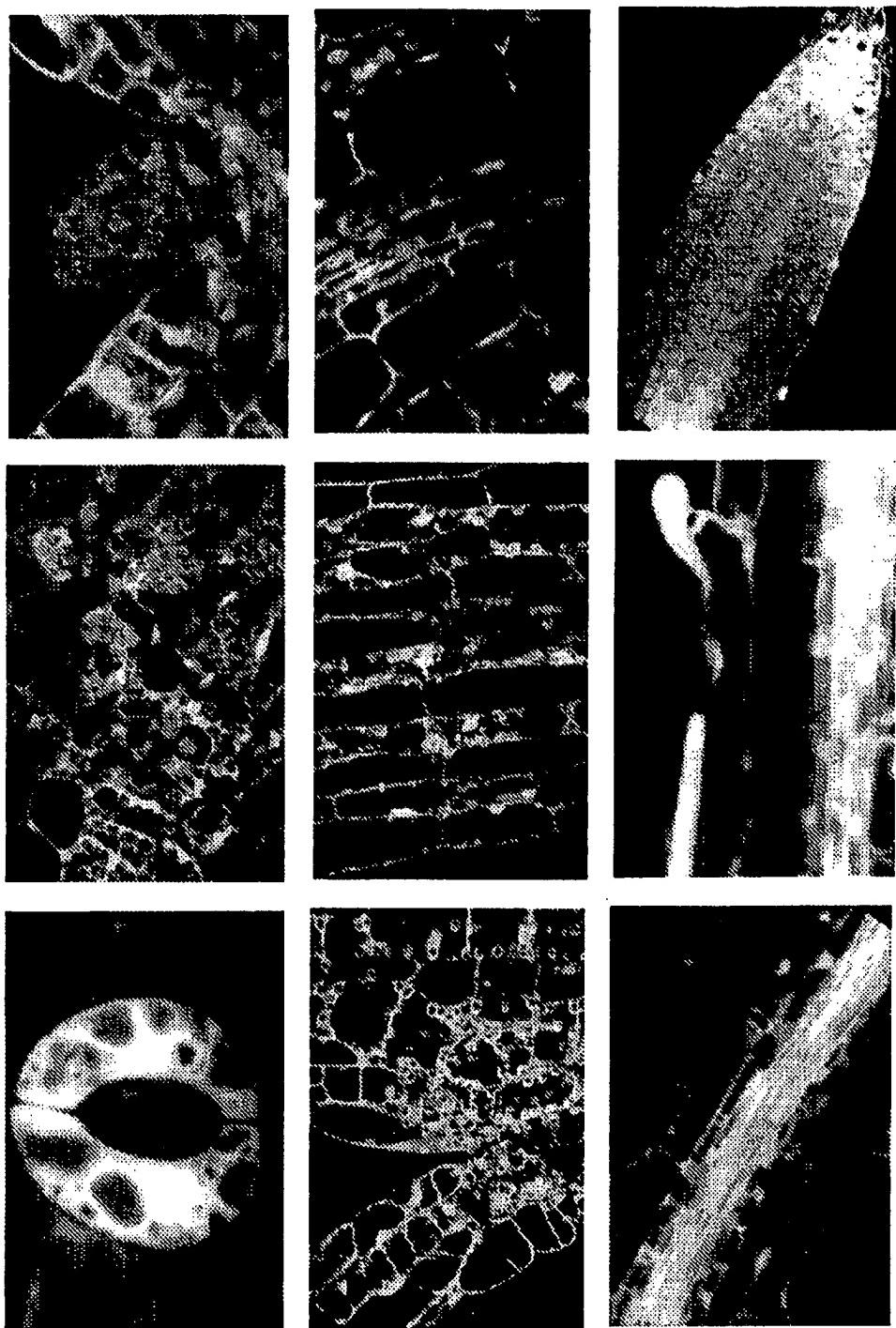
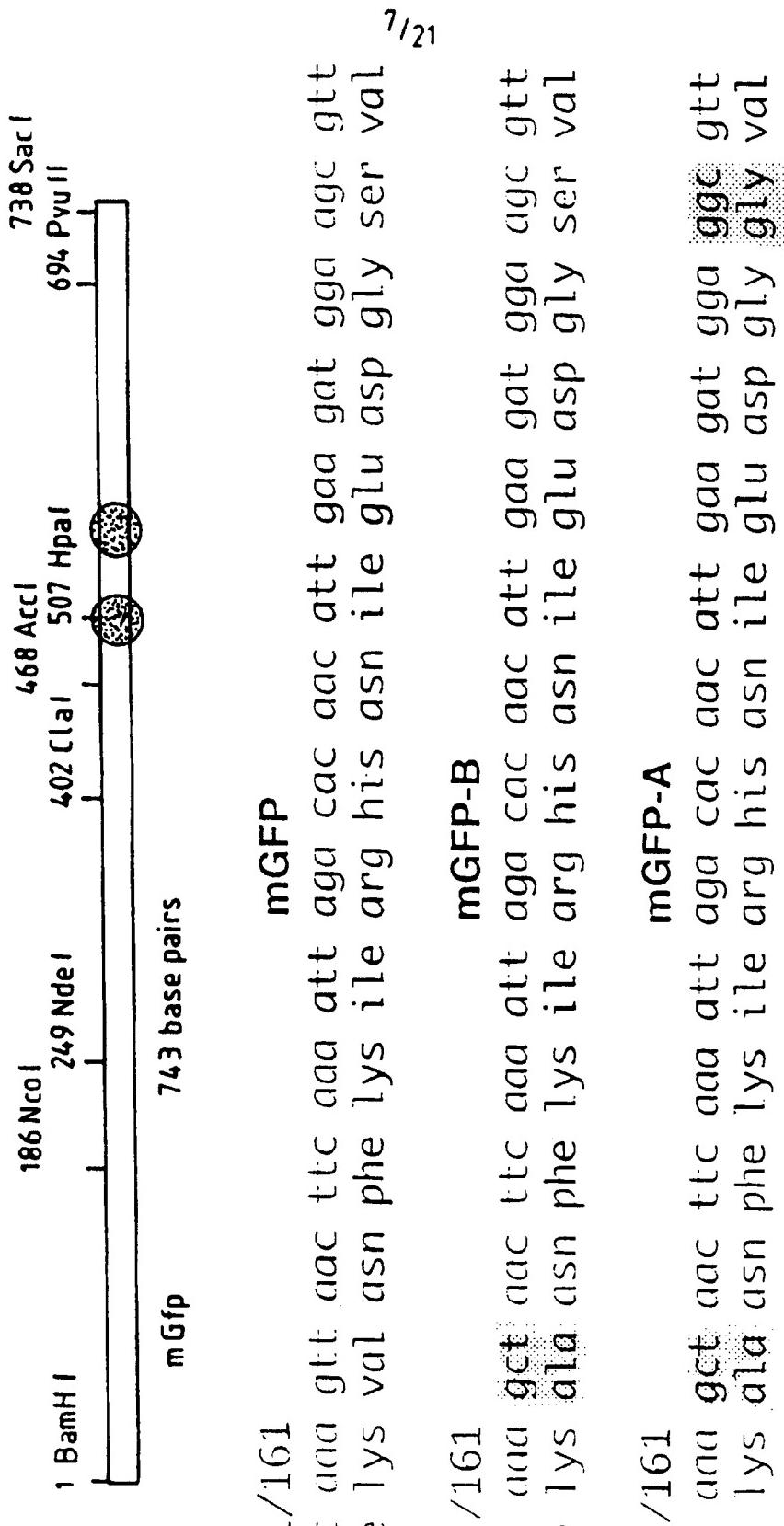


Fig. 5



SUBSTITUTE SHEET (RULE 26)

Fig. 6

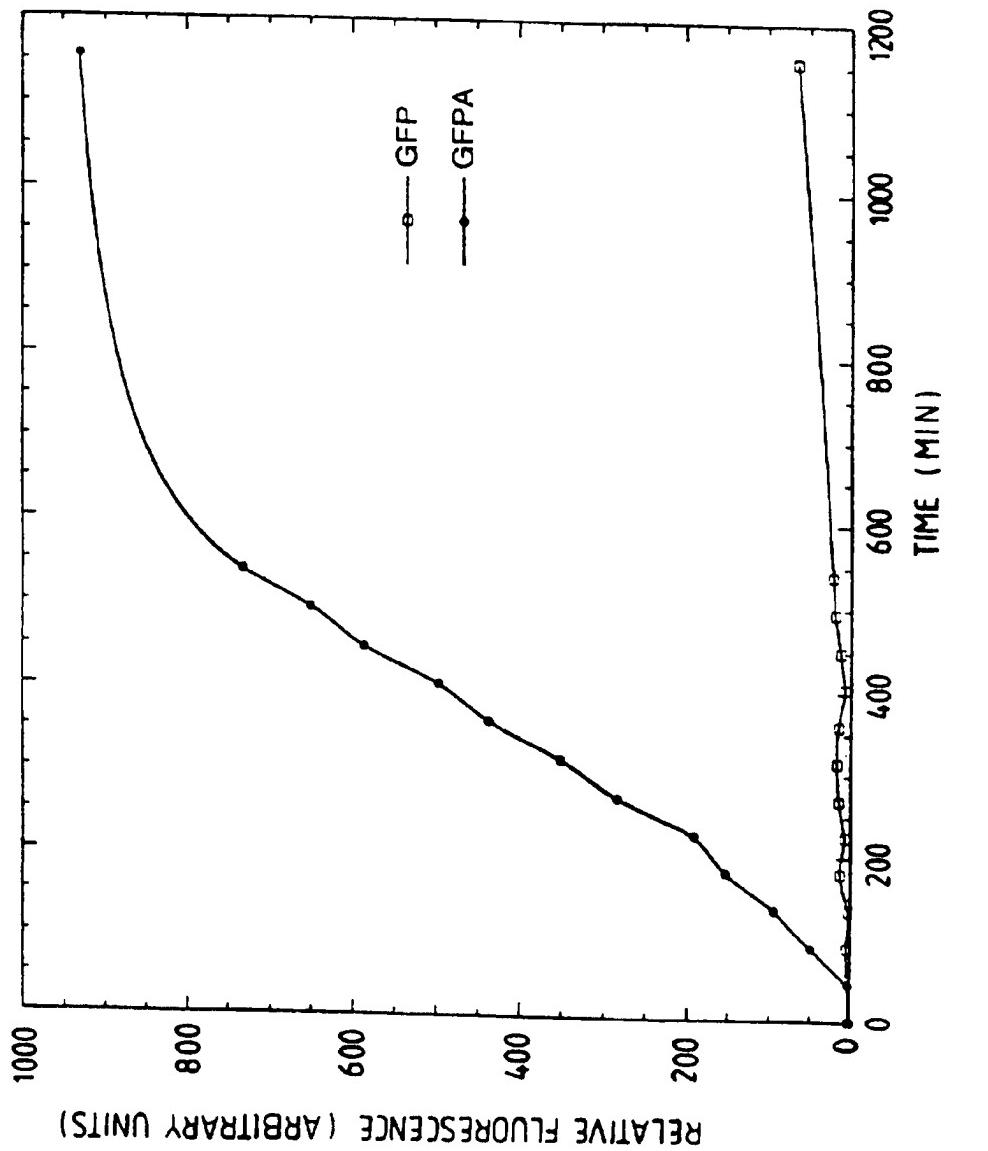
θ_{21} 

Fig. 7

9/21

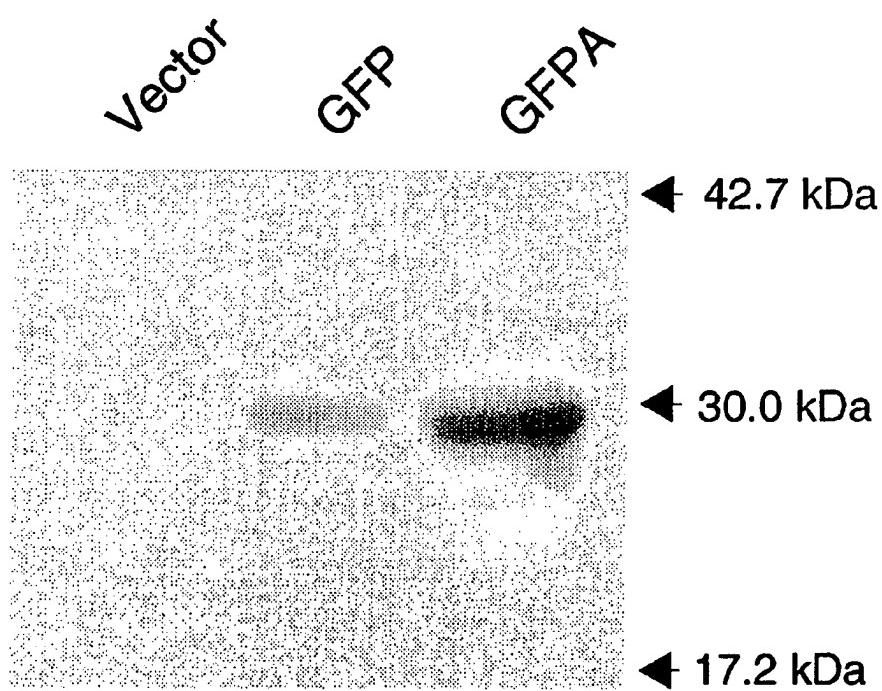
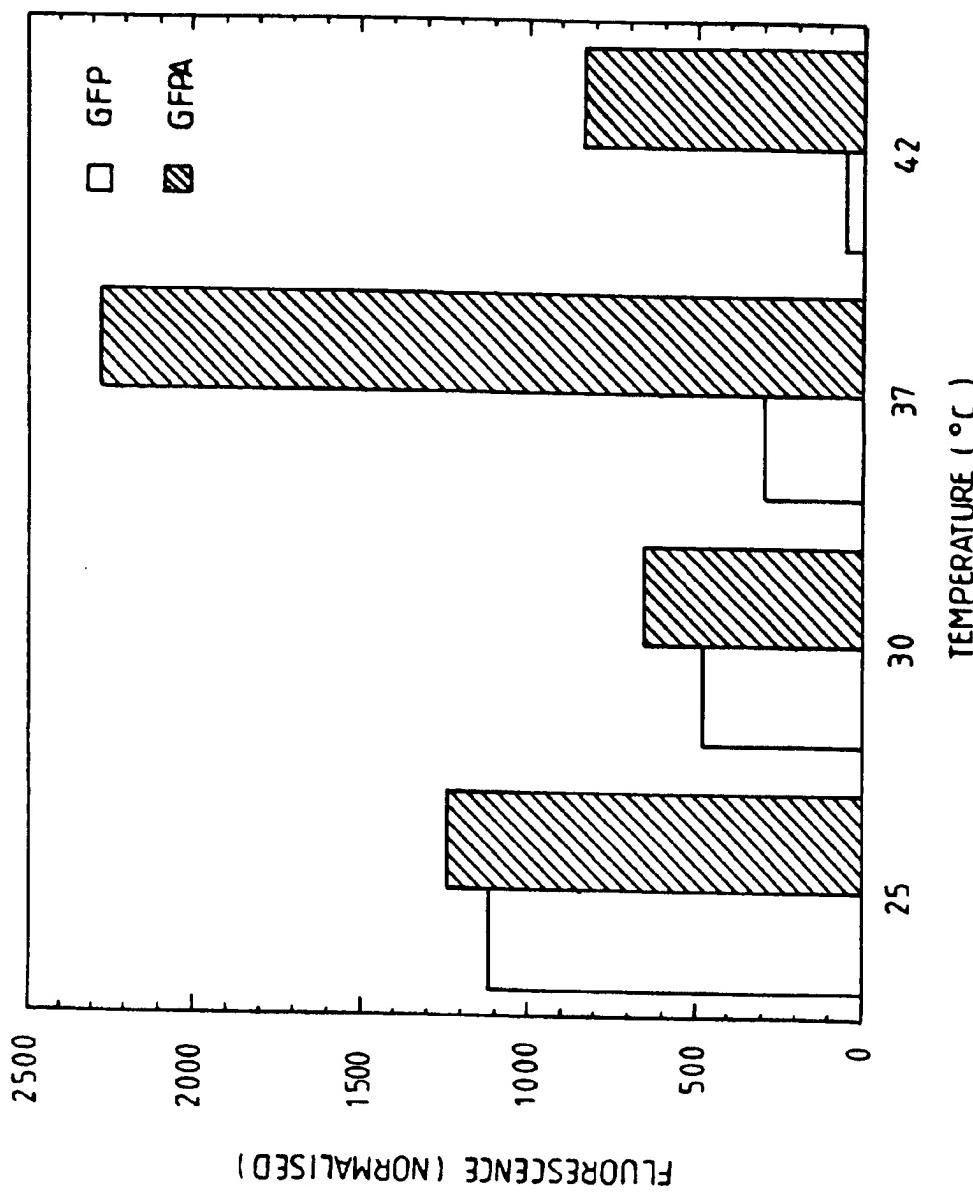


Fig. 8

10/21

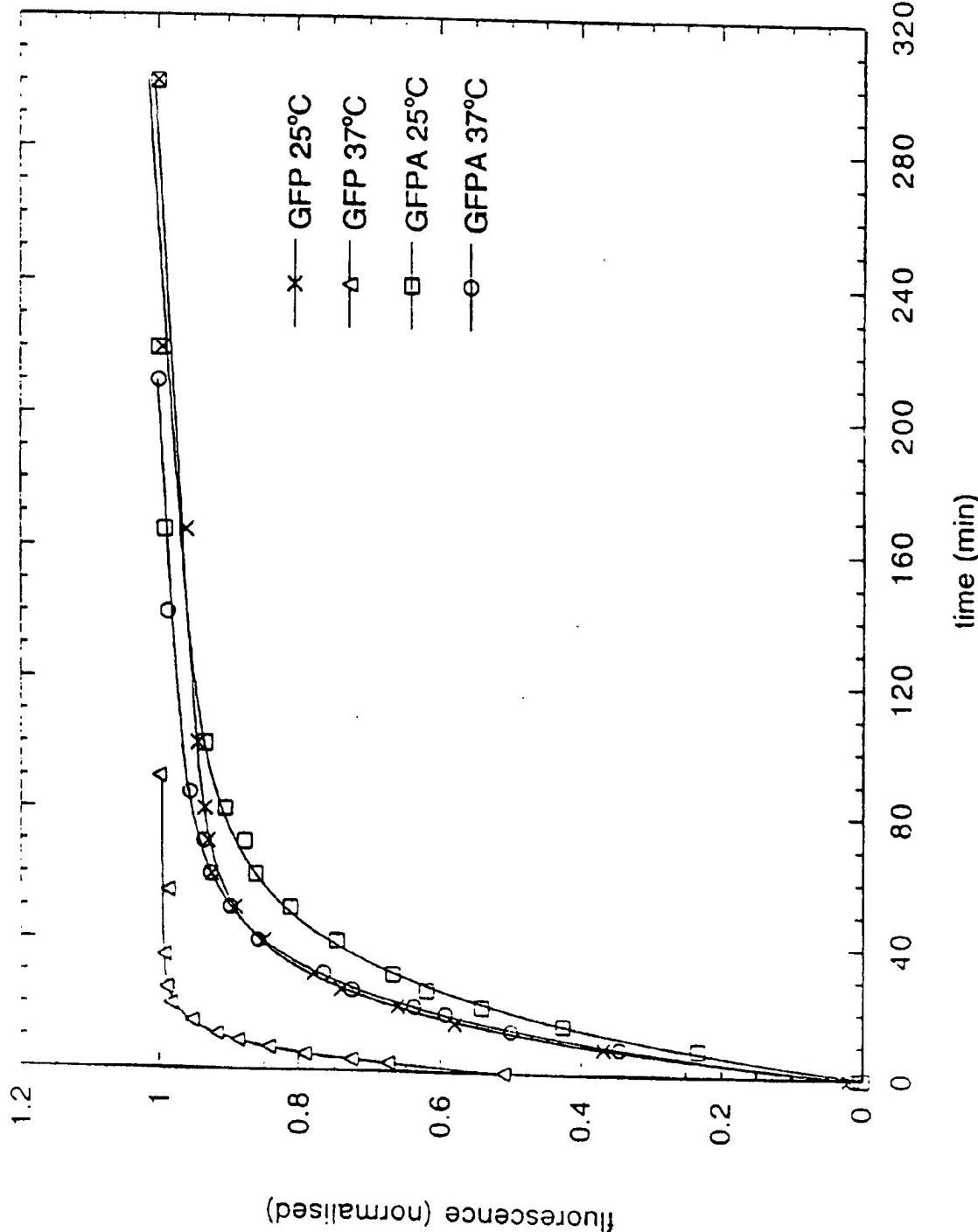
Fig. 9



SUBSTITUTE SHEET (RULE 26)

11/21

Fig. 10



12/21

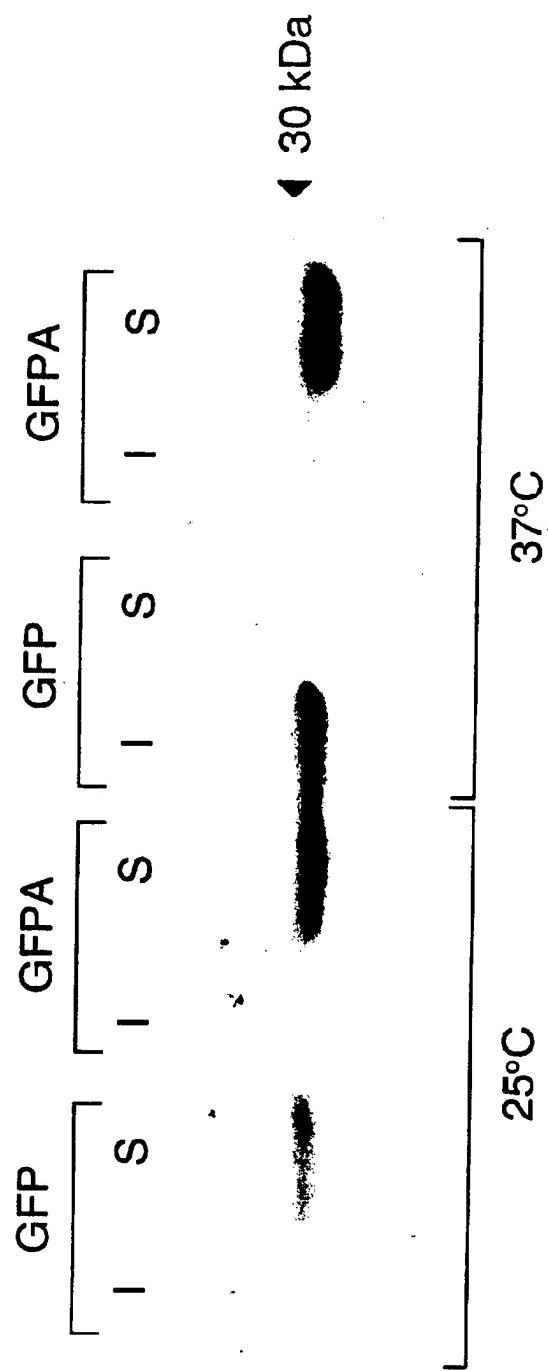
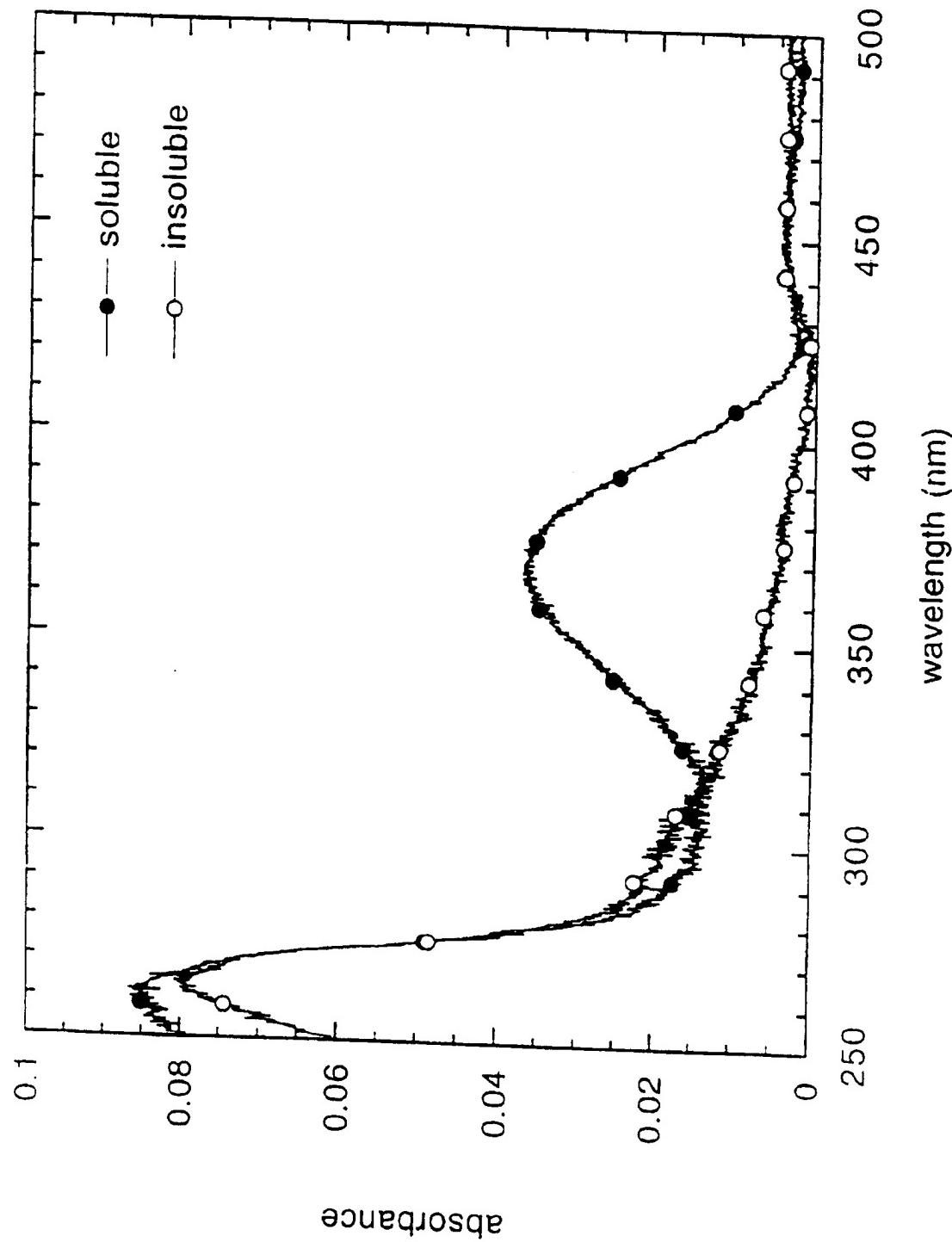


Fig. 11

¹³/21

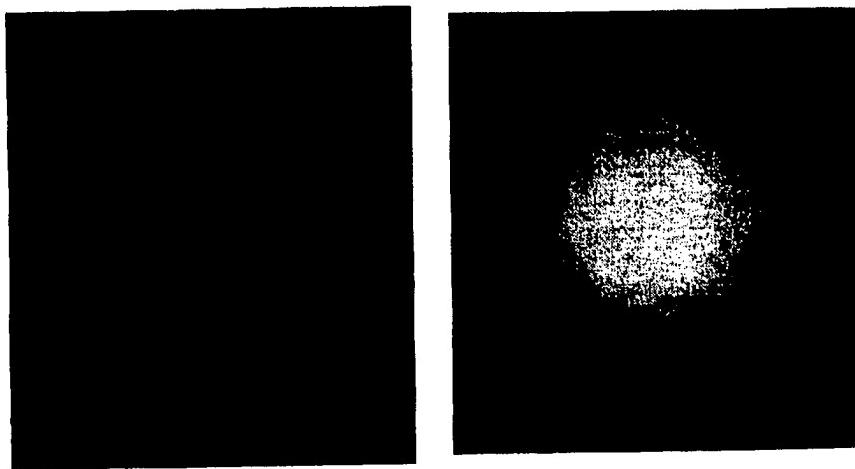
Fig. 12



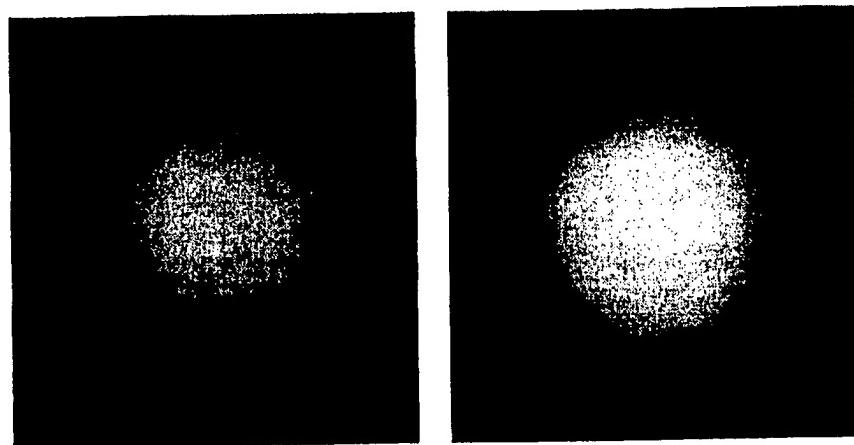
SUBSTITUTE SHEET (RULE 26)

14/21

37°C



25°C



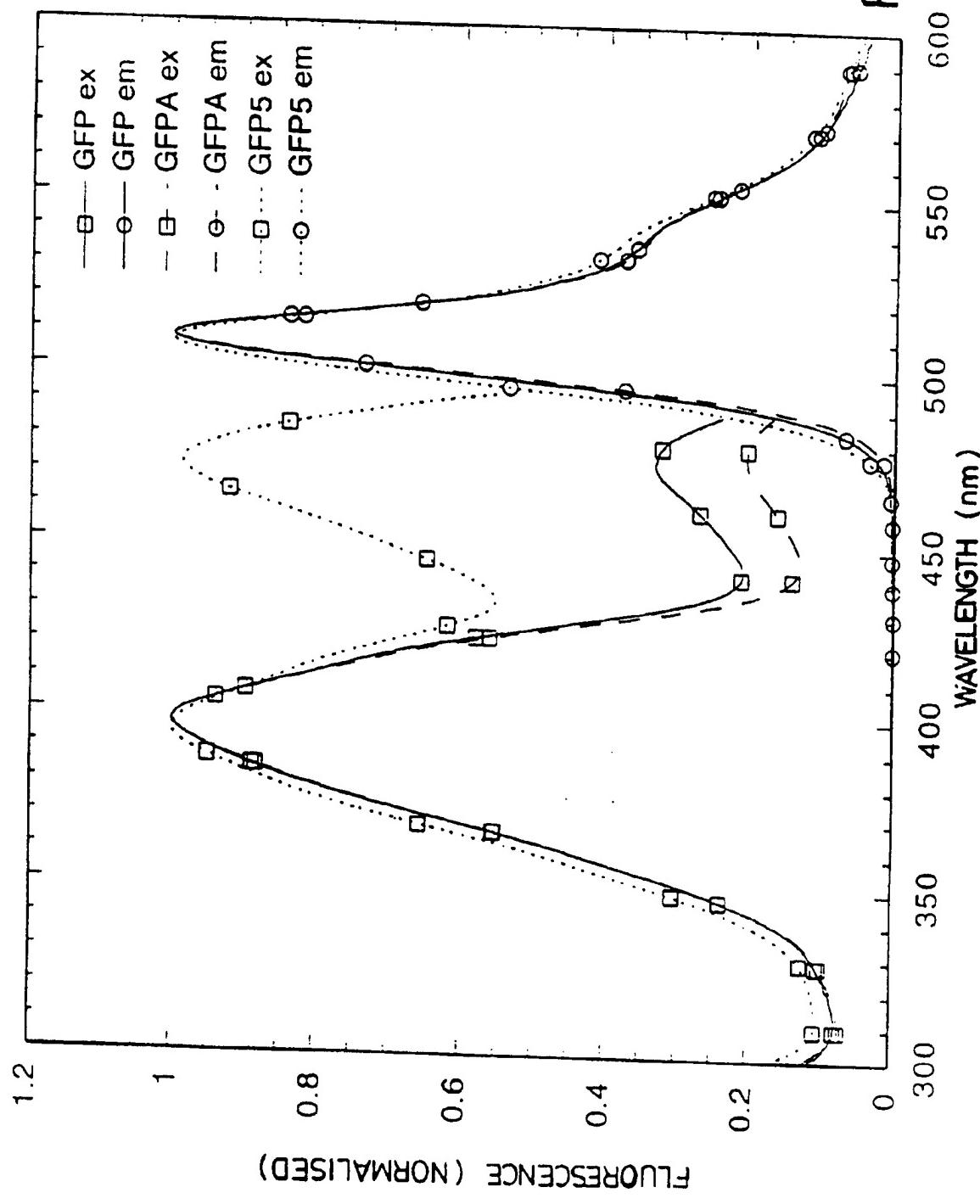
GFP

GFP-A

Fig. 13

15/21

Fig. 14



16/21

etc act eaa gga cca gaa ctt ttc act gga ctt gtc cca att ccc gtt gaa tra gat
 atg act aaa gga gaa ctt ttc act gga gtt gtc cca att ccc gtt gaa tra gat
 M S K G E L F T G V P I L V E L D

ggg gat gtt aat ggg cac aaa ttt tct gtc aat gga gag ggt gaa gat gca aca tac
 ggt gat gtt aat ggg ccc aat ttt tct gtc aat gga gag ggt gaa gat gca aca tac
 C D V N G H K F S V S G E G D A T Y

gga aaa ctt acc ctt aaa ttt att tgc act act gga aaa cta cct gtt cca tgg cca aca
 gga aaa ctt acc ctt aaa ttt att tgc act act gga aaa cta cct gtt cca tgg cca aca
 G K L T L K F I C T G K P V P W P T

Nco I

ctt gtc act act ttc tct tat ggt gtt caa tgc ttt tca aga tac cca gat cat atg aat
 ctt gtc act act ttc tct tat ggt gtt caa tgc ttt tca aga tac cca gat cat atg aat
 L V T T F S Y G V Q C F S R Y P D H M K

Nde I

cgg cat gac ttt ttc aag aat gcc atg ccc gaa ggt tat gta gag age act ata ttt
 cgg cac gac ttc ttc aat aat gcc atg ccc gaa ggt tat gta gag age act ata ttt
 R H D F K S A M P E G Y V Q E R T I F

tcc aaa gat gac ggg aac tac aag aca cgt gct gaa gtc aag ttt gaa ggt gat acc ctt
 ttc aat gac gac ggg aac tac aag aca cgt gct gaa gtc aag ttt gaa ggt gat acc ctt
 F K D D G N Y K T R A E V K F E G D T L

cryptic intron
 gtt eat aat gac atc gag tte eaa gat gtt gat ttt aaa gaa gat cga aac att ctt gga cac
 gtc eac aat gac atc gag ctt aag gga atc gat ttc aag gag gac gga aac atc ctc ggc cac
 V N E I E L K G I D F K E D G N I L C H

SUBSTITUTE SHEET (RULE 26)

Fig.15 Sheet 1

17/21

<u>Acc I</u>											
aaa ttg aea tac aac tac aac tca cac eat qta tac atc atq qca gac aat											gfp m-gfps
aag ttg gaa tac aac tac eat tcc cac aac gta tac atc atq gcc gac aag aac											m-gfps
K L E Y N S H N V I M A D K Q K N											
<u>V</u>											
gga atc aaa gtt aac ttc aaa att aga cac aac att gaa gat gga agc gtt caa cta gca											gfp m-gfps
ggc etc aaa gcc eac ttc aag acc cgc cac aac atc gaa gac ggc ggc gtg caa ctc gct	I	K	A	N	F	T	R	H	N	I	E D G G V Q L A
D H Y Q N T P I G D G P V L L P D N H											
<u>I</u>											
gac cat tat caa caa aat act cca att ggc gat ggc cct gtc ctt tta cca gac aac cat											gfp m-gfps
gat cat tat caa caa aat act cca att ggc gat ggc cct gtc ctt tta cca gac aac cat	D	H	Y	Q	N	T	P	I	G	D	G V L P D N H
<u>S</u>											
tac ctg tcc aca caa tcc gcc ctt tcg aca gat ccc aac gaa aag aca gac cac atg gtc											gfp m-gfps
tac ctg tcc aca caa tcc gcc ctt tcg aca gat ccc aac gaa aag aca gac cac atg gtc	Y	L	S	T	Q	S	A	L	S	K	D P N E K R D H M V
<u>E</u>											
ctt ctt gag ttG Gta aca gct gct ggg att aca cat gtc atg gat gaa cta tac aat taa											gfp m-gfps
ctt ctt gag ttG Gta aca gct gct ggg att aca cat gtc atg gat gaa cta tac aat taa	L	E	F	V	T	A	A	G	I	T H G M D E L Y K	*

SUBSTITUTE SHEET (RULE 26)

Fig. 15 Sheet 2

18/21

21/1	ATG	AAG	ACT	AAT	C'TT	TTT	CTC	TTC	ATC	TTT	TCA	CTT	CTC	CTA	TCA	TCA	TCC	TCA	TCG	GCC
M	K	T	N	L	F	L	F	L	I	F	S	L	L	S	I	S	I	S	S	A
81/21	GAA	TTC	agt	aaa	ggg	gaa	gaa	ctt	ttc	act	ggg	gtt	gtc	cca	att	ctt	gtt	gaa	tta	gtt
E	F	S	K	G	E	E	E	L	F	T	G	V	V	P	I	L	V	E	L	D
141/41	ggg	gtt	aat	ggg	cac	aaa	ttt	tct	gtc	agt	ggg	gag	ggt	gaa	ggg	gtt	gca	aca	tac	
G	D	V	N	G	H	K	F	S	V	S	G	E	G	E	D	A	T	Y		
201/61	ggg	aaa	ctt	acc	ctt	aaa	ttt	att	tgc	act	act	ggg	aaa	cta	cct	gtt	cca	tgg	cca	aca
G	K	L	T	L	K	F	I	C	T	T	G	K	L	P	V	P	W	P	T	
261/81	ctt	gtc	act	act	ttc	tct	tat	ggt	gtt	caa	tgc	ttt	tca	aga	tac	cca	gat	cat	atg	aag
L	V	T	T	F	S	Y	G	V	Q	C	F	S	R	Y	P	D	H	M	K	
321/101	cgg	cac	gac	tcc	tcc	aag	agc	gcc	atg	cct	gag	gga	tac	gtg	cag	gag	agg	acc	atc	ttc
R	H	D	F	F	K	S	A	M	P	E	G	Y	V	Q	E	R	T	I	F	
381/121	ttc	aag	gac	gac	ggg	aac	tac	aag	aca	cgt	gct	gaa	gtc	aag	ttt	gag	gga	gac	acc	
F	K	D	G	N	Y	K	T	R	A	E	V	K	F	E	G	D	T	L		
411/131	411/131	411/131	411/131	411/131	411/131	411/131	411/131	411/131	411/131	411/131	411/131	411/131	411/131	411/131	411/131	411/131	411/131	411/131	411/131	

SUBSTITUTE SHEET (RULE 26)**Fig. 16 Sheet 1**

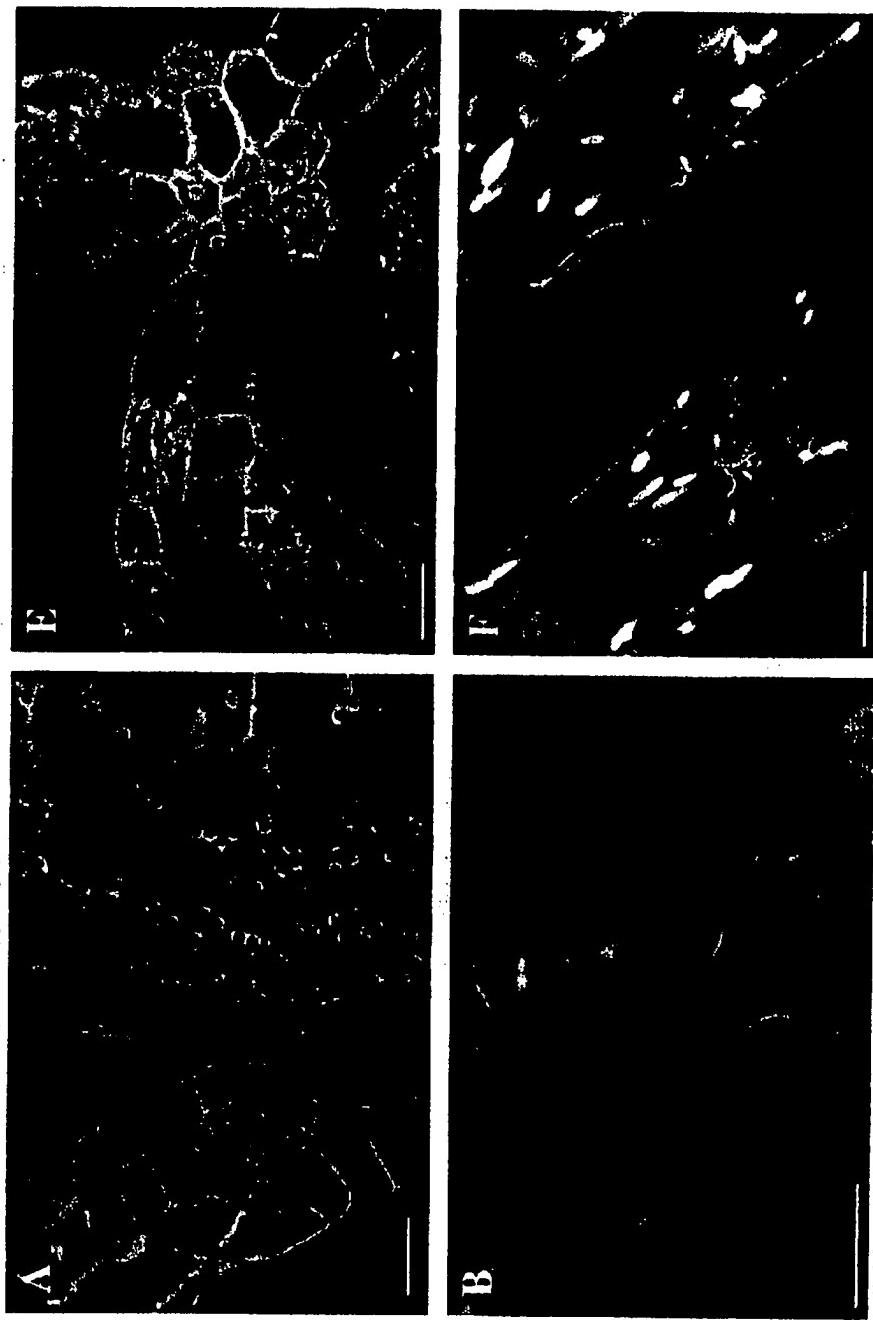
19/21

441/141	gtc aac agg atc gag ctt aag gga atc gat	ttc aag gag gac gga aac atc ctc gga cac	
V N R I E L K G T D F K E D G N I L G H	501/161	aag tlg gaa tac aac tcc cac aac gta tac atc atg gcc gac aag cag aag aac	
K I, E Y N Y N S H N V Y I M A D K Q K N	561/181	ggc atc aaa gcc aac ttc aag acc cgc cac aac atc gaa gac ggc gtc gaa ctc gct	
G I K A N F K T R H N I E D G V Q L A	621/201	651/211	gat cat tat caa caa aat act cca att ggc gat ggc ctc ctt tta cca gac aac cat
D H Y Q N T P I G D G P V L L P D N H	681/221	tac ctg tcc aca caa tct gcc ctt tcg aaa gat ccc aac gaa aag aga gac cac atg gtc	
Y L S T Q S A L S K D P N E K R D H M V	741/241	711/231	ctt ctt gag ttt gta aca gct ggg att aca cat ggc atg gat gaa cta tac aaa cac
L L E F V T A G I T H G M D E L Y K H	801/261	gac gaa ctc taa	
D E L *			

Fig. 16 Sheet 2

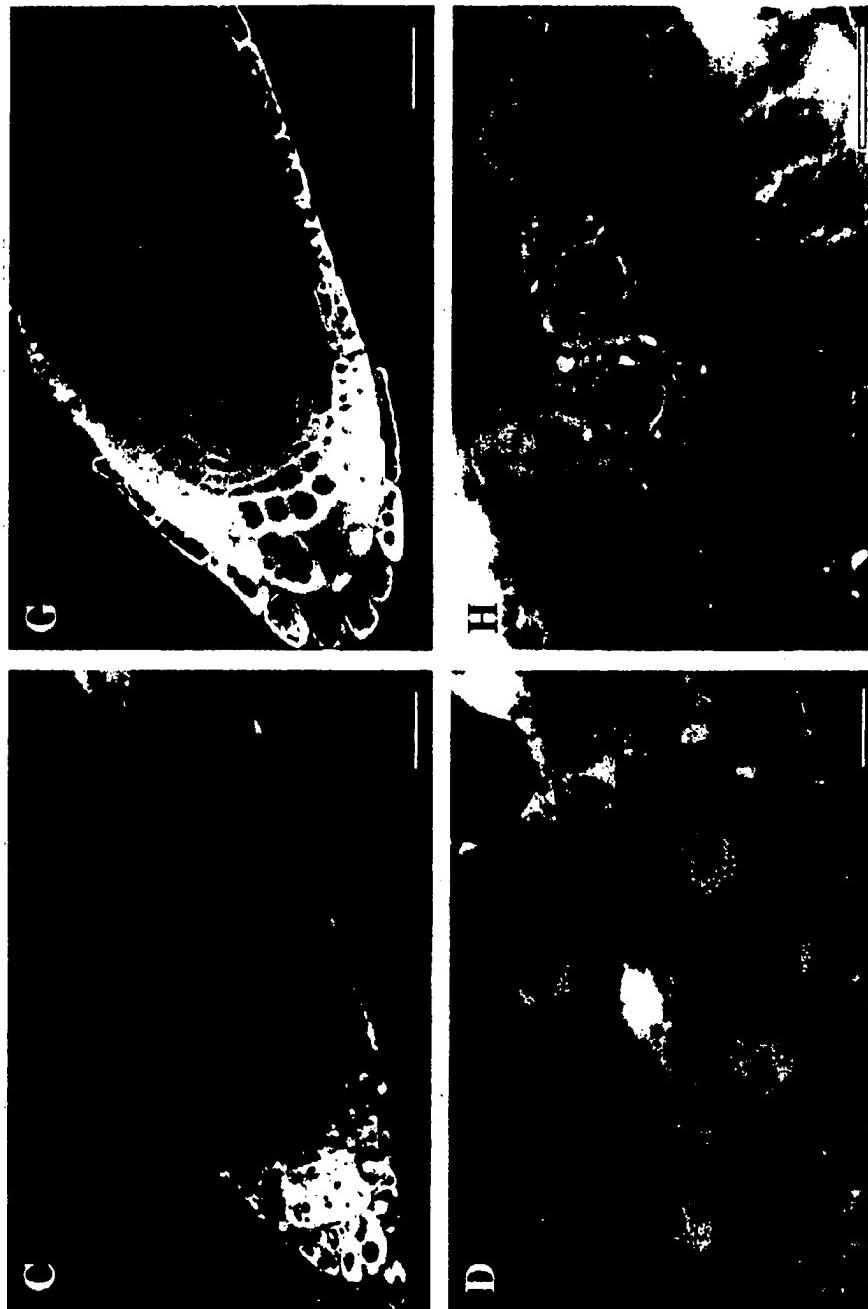
20/21

Fig. 17 Sheet 1



2/21

Fig. 17 Sheet 2



INTERNATIONAL SEARCH REPORT

Int. Search Report
PCT/GB 94/0481

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/82 C12N15/12 C07K14/435 G01N33/52 G01N33/53

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	TRENDS IN GENETICS, vol. 11, no. 8, August 1995, pages 328-329, XP002003595 HASELOFF, J., ET AL.: "GFP in plants" see the whole document ---	1-5,10, 20,21,24
P,X	CURR. BIOL. (1996), 6(3), 325, pages 325-330, XP000571865 CHUI, W., ET AL.: "Engineered GFP as a vital reporter plants" see the whole document ---	1-5,10, 20,21,24
P,X	WO,A,95 07463 (UNIV COLUMBIA ;WOODS HOLE OCEANOGRAPHIC INST (US); CHALFIE MARTIN) 16 March 1995 see the whole document ---	1 -/-

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

* Special categories of cited documents :

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubt on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

& document member of the same patent family

2

Date of the actual completion of the international search 28 May 1996	Date of mailing of the international search report 07.06.96
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patenttaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax (+ 31-70) 340-3016	Authorized officer Maddox, A

Form PCT/ISA-210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 96/00481

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	PLANT CELL REPORTS, vol. 14, April 1995, pages 403-406, XP000571886 NIEDZ, R.P., ET AL.: "Green fluorescent protein : an in vivo reporter of plant gene expression" see the whole document ---	23,24
P,X	THE PLANT JOURNAL, vol. 8, no. 5, November 1995, pages 777-784, XP002003596 SHEEN, J., ET AL.: "Green-fluorescent protein as a new vital marker in plant cells" see the whole document ---	23,24
A	SCIENCE, vol. 263, 11 February 1994, pages 802-805, XP002003599 CHLAFIE, M., ET AL.: "Green fluorescent protein as a marker for gene expression" see the whole document ---	1-23
A	WO,A,91 01305 (UNIV WALES MEDICINE) 7 February 1991 see claim 10 ---	1-23
A	NATURE, vol. 369, June 1994, pages 400-403, XP002003600 WANG, S., ET AL.: "Implications for bcd mRNA localization from spatial distribution of exu protein in Drosophila oogenesis" see the whole document -----	23

INTERNATIONAL SEARCH REPORT

International Application No.

PCT/GB 95/01481

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO-A-9507463	16-03-95	US-A-	5491084	13-02-96
		AU-B-	7795794	27-03-95
		CA-A-	2169298	16-03-95
-----	-----	-----	-----	-----
WO-A-9101305	07-02-91	AU-B-	6054590	22-02-91
		EP-A-	0484369	13-05-92
		JP-T-	5501862	08-04-93
-----	-----	-----	-----	-----

(1993 97 CWR FCL)

PCT

ORGANISATION MONDIALE DE LA PROPRIETE INTELLECTUELLE
Bureau international



DEMANDE INTERNATIONALE PUBLIEE EN VERTU DU TRAITE DE COOPERATION EN MATIERE DE BREVETS (PCT)

(51) Classification internationale des brevets ⁶ : G01N 33/58		A2	(11) Numéro de publication internationale: WO 98/55873 (43) Date de publication internationale: 10 décembre 1998 (10.12.98)
(21) Numéro de la demande internationale: PCT/FR98/01136 (22) Date de dépôt international: 4 juin 1998 (04.06.98)		(81) Etats désignés: CA, JP, US, brevet européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(30) Données relatives à la priorité: 97/06977 5 juin 1997 (05.06.97) FR		Publiée <i>Sans rapport de recherche internationale, sera republiée dès réception de ce rapport.</i>	
(71) Déposant (<i>pour tous les Etats désignés sauf US</i>): CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE [FR/FR]; 3, rue Michel-Ange, F-75794 Paris Cedex 16 (FR).			
(72) Inventeurs; et			
(75) Inventeurs/Déposants (<i>US seulement</i>): GALZI, Jean-Luc [FR/FR]; 37, rue Saint Aloise, F-67100 Strasbourg (FR). ALIX, Philippe [FR/FR]; 11, rue de la Gare, F-14650 Carpiquet (FR).			
(74) Mandataires: GROSSET-FOURNIER, Chantal etc.; Grosset-Fournier & Demachy S.A.R.L., 103, rue La Fayette, F-75481 Paris Cedex 10 (FR).			

(54) Title: USE OF A FLUORESCENT PROTEIN FOR DETECTING INTERACTION BETWEEN A TARGET PROTEIN AND ITS LIGAND

(54) Titre: UTILISATION D'UNE PROTEINE FLUORESCENTE POUR LA DETECTION D'INTERACTIONS ENTRE UNE PROTEINE CIBLE ET SON LIGAND

(57) Abstract

The invention concerns the use of a fluorescent protein selected in particular among the autofluorescent proteins for detecting the non-covalent interaction between a target protein marked by the fluorescent protein and one of its ligands marked by a marker consisting of: either a molecule capable of absorbing the light emitted by the fluorescent protein, or a fluorescent substance, said detection taking place by fluorescence energy transfer: between the fluorescent protein and said fluorescent substance, the fluorescent substance being such that it is excitable at the fluorescent protein emitting wavelength, or it emits at the fluorescent protein emitting wavelength; between the fluorescent protein and said molecule capable of absorbing the light emitted by the fluorescent protein.

(57) Abrégé

L'invention concerne l'utilisation d'une protéine fluorescente choisie notamment parmi les protéines autofluorescentes, pour la détection de l'interaction non covalente entre une protéine cible marquée par la protéine fluorescente et l'un de ses ligands marqué par un marqueur constitué: soit par une molécule susceptible d'absorber la lumière émise par la protéine fluorescente, soit par une substance fluorescente, cette détection ayant lieu par transfert d'énergie de fluorescence: entre la protéine fluorescente et la susdite substance fluorescente, la substance fluorescente étant telle que soit elle est excitable à la longueur d'onde d'émission de la protéine fluorescente, soit elle émet à la longueur d'excitation de la protéine fluorescente, ou entre la protéine fluorescente et la susdite molécule susceptible d'absorber la lumière émise par la protéine fluorescente.

UNIQUEMENT A TITRE D'INFORMATION

Codes utilisés pour identifier les Etats parties au PCT, sur les pages de couverture des brochures publant des demandes internationales en vertu du PCT.

AL	Albanie	ES	Espagne	LS	Lesotho	SI	Slovénie
AM	Arménie	FI	Finlande	LT	Lituanie	SK	Slovaquie
AT	Autriche	FR	France	LU	Luxembourg	SN	Sénégal
AU	Australie	GA	Gabon	LV	Lettonie	SZ	Swaziland
AZ	Azerbaïdjan	GB	Royaume-Uni	MC	Monaco	TD	Tchad
BA	Bosnie-Herzégovine	GE	Géorgie	MD	République de Moldova	TG	Togo
BB	Barbade	GH	Ghana	MG	Madagascar	TJ	Tadjikistan
BE	Belgique	GN	Guinée	MK	Ex-République yougoslave de Macédoine	TM	Turkménistan
BF	Burkina Faso	GR	Grèce	ML	Mali	TR	Turquie
BG	Bulgarie	HU	Hongrie	MN	Mongolie	TT	Trinité-et-Tobago
BJ	Bénin	IE	Irlande	MR	Mauritanie	UA	Ukraine
BR	Brésil	IL	Israël	MW	Malawi	UG	Ouganda
BY	Bélarus	IS	Islande	MX	Mexique	US	Etats-Unis d'Amérique
CA	Canada	IT	Italie	NE	Niger	UZ	Ouzbékistan
CF	République centrafricaine	JP	Japon	NL	Pays-Bas	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norvège	YU	Yougoslavie
CH	Suisse	KG	Kirghizistan	NZ	Nouvelle-Zélande	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	République populaire démocratique de Corée	PL	Pologne		
CM	Cameroun	KR	République de Corée	PT	Portugal		
CN	Chine	KZ	Kazakhstan	RO	Roumanie		
CU	Cuba	LC	Sainte-Lucie	RU	Fédération de Russie		
CZ	République tchèque	LI	Liechtenstein	SD	Soudan		
DE	Allemagne	LK	Sri Lanka	SE	Suède		
DK	Danemark	LR	Libéria	SG	Singapour		
EE	Estonie						

UTILISATION D'UNE PROTEINE FLUORESCENTE POUR LA DETECTION D'INTERACTIONS ENTRE UNE PROTEINE CIBLE ET SON LIGAND.

5

L'invention concerne l'utilisation d'une protéine fluorescente pour la détection d'interactions entre une protéine cible et son ligand.

De nombreux médicaments et substances naturelles exercent leur action en interagissant avec des protéines régulatrices appelées récepteurs, impliquées dans de nombreuses fonctions physiologiques des organismes, et les altérations de leur fonctions sont la cause de nombreuses pathologies. L'accessibilité des récepteurs aux agents pharmacologiques endogènes ou exogènes, naturels ou synthétiques, depuis l'extérieur de la cellule conduit à les considérer comme des cibles de choix pour la recherche de molécules biologiquement actives, en particulier de molécules présentant des pouvoirs thérapeutiques potentiels.

Afin d'identifier de nouveaux outils pharmacologiques et produits médicamenteux, plusieurs tests de criblage de molécules biologiquement actives, ont été développés :

20 - le test SPA ou test de proximité par scintillation (scintillation proximity assay) (Udenfirend *et al.* Anal. Biochem. (1987) 161:494-500, brevet américain n° 4,568,649; brevet européen 0,154,734 et demande de brevet japonais n° 84//52452) ne parvient pas à s'affranchir de l'utilisation de molécules radioactives, et donc de toutes les nuisances associées à la manipulation, l'utilisation et le stockage de radioéléments;

25 - le test de mesure par transfert d'énergie de fluorescence utilisant le couple europium - allophycocyanine (Mathis. clin. Chem. (1993), 39:1953-1959) requiert l'utilisation d'allophycocyanine purifiée et nécessite son greffage sur une protéine cible elle aussi purifiée, ce qui n'est pas toujours possible surtout lorsque la protéine cible n'est pas abondante ;

30 - les tests fonctionnels faisant intervenir la régulation de gènes rapporteurs codant pour des protéines luminescentes, ou pourvues d'une activité enzymatique dosée par colorimétrie, sont des mesures très indirectes de l'interaction entre un ligand et son récepteur, peuvent être source de faux-positifs, font intervenir des cascades d'amplification qui interfèrent avec la mesure quantitative des interactions entre le récepteur et ses ligands, et ne sont applicables que sur des système biologiques couplés à une transcription des gènes (Broach, J.R. & Thorner, J. 1996, Nature 384 supp. 14-16) ;

- un test de mesure par fluorescence permettant la détection de la liaison d'un messager secondaire intracellulaire (le nucléotide cyclique AMPc) développé par greffage d'un site de liaison du messager secondaire sur la protéine fluorescente verte (GFP) de méduse *Aequorea victoria* (Thastrup *et al.* WO96/23898). Ce test qui permet la détection de molécules biologiquement actives se limite aux analogues de messagers secondaires affectant des processus intracellulaires.

Le gène codant pour une protéine fluorescente issue de la méduse *Aequorea victoria*, la protéine fluorescente verte (green fluorescent protein ou GFP) (Prasher *et al.* 1992, Gene 111:229-233), a été récemment déchiffré. La GFP est une protéine monomérique. Elle acquiert ses propriétés de fluorescence par un mécanisme autocatalytique de formation du fluorophore. L'expression, endogène ou hétérologue, de la GFP ne requiert qu'un gène et ne nécessite l'addition d'aucun groupe prosthétique. La GFP a été exprimée de manière hétérologue dans des cellules et organismes aussi divers que les bactéries, les levures, les cellules eucaryotes animales et végétales. La structure de la GFP (Ormö *et al.* 1996, Science 273:1392-1395 ; Yang *et al.* 1996, Nature Biotechnology, 14:1246-1251) permet son greffage sur d'autres polypeptides, soit du côté amino-terminal soit du côté carboxy terminal sans préjudice ni vis-à-vis de son taux d'expression, ni vis-à-vis de la formation du fluorophore. La GFP a ainsi pu être couplée, par fusion de gènes, à des protéines solubles ou membranaires. Enfin, les codons contenus dans le gène naturel ont été remplacés par les codons préférés des organismes hôtes (EGFP [Cormack *et al.* 1996, Gene 173:33-38] par exemple, pour l'expression dans les cellules eucaryotes animales) et diverses mutations permettent de modifier ses spectres d'absorption et d'émission de lumière et rendent possible des détections multiples dans un système d'expression unique.

L'invention a pour but notamment la préparation de protéines cibles, notamment des récepteurs, rendu(e)s fluorescent(e)s par fusion avec une protéine fluorescente, d'une part, de leurs ligands marqués, d'autre part, et leur utilisation pour la détection d'interactions entre les récepteurs fluorescents et leurs ligands marqués et l'identification de nouvelles molécules biologiquement actives.

L'un des autres aspects de l'invention est un procédé simple à mettre en oeuvre, rapide, sensible, permettant d'effectuer des mesures quantitatives, à l'équilibre et en temps réel, d'interactions non covalentes entre une protéine cible et son ligand.

L'un des autres aspects de l'invention est de proposer un procédé généralisable à de nombreuses protéines cibles et à leurs ligands.

L'un des autres aspects de l'invention est de proposer un procédé ne nécessitant ni la purification de la protéine cible, ni celle du ligand.

5 L'un des aspects de l'invention est de proposer un procédé non polluant puisqu'il n'utilise pas de radioactivité, économique puisqu'il utilise la lumière visible (pas de quartz) sur les équipements disponibles et ne nécessitant aucune filtration.

L'un des autres aspects est de proposer un procédé automatisable.

10 L'invention a pour objet la construction de d'ADNc contenant une séquence codant pour

15 1) une protéine, notamment un récepteur, dans lequel un ou plusieurs acides aminés sont substitués, insérés ou déletés, fusionnée avec l'ADNc codant pour la GFP, ou l'un de ses mutants, dans lequel un ou plusieurs acides aminés sont substitués, insérés ou déletés afin de maintenir la phase de lecture et d'obtenir un polypeptide hybride.

20 L'invention a également pour objet des cellules contenant une séquence d'ADN codant pour une protéine, notamment un récepteur, dans lequel un ou plusieurs acides aminés sont substitués, insérés ou déletés, fusionnée avec l'ADNc codant pour la GFP, ou l'un de ses mutants, dans lequel un ou plusieurs acides aminés sont substitués, insérés ou déletés afin de maintenir la phase de lecture et capable d'exprimer la séquence d'ADN considérée.

25 L'invention a également pour objet la production de ligands pharmacologiquement actifs contenant un groupe chimique fluorescent lié par une réaction chimique dans lequel le groupe fluorescent est soit donneur d'énergie pour la GFP ou l'un de ses mutants, soit accepteur d'énergie de la GFP ou de l'un de ses mutants.

30 L'invention a également pour objet la culture de cellules contenant l'hybride protéine-GFP dans des conditions permettant

- 30 1) l'expression du polypeptide hybride,
- 2) la détection de fluorescence de la cellule.

35 L'invention a également pour objet l'incubation de cellules exprimant l'hybride protéine-GFP avec le ligand fluorescent, la mesure des changements de fluorescence soit d'émission du donneur, soit d'émission de l'accepteur, soit d'excitation de l'accepteur, révélant l'interaction protéine ligand, ainsi que l'addition d'une molécule suspectée biologiquement active avec le ligand fluorescent et la mesure des altérations du signal de transfert d'énergie par rapport à l'incubation des susdites cellules avec le ligand fluorescent.

L'invention a également pour objet la construction d'ADNc contenant une séquence codant pour

1) un polypeptide, notamment un ligand, dans lequel un ou plusieurs acides aminés sont substitués, insérés ou délétés, fusionné avec l'ADNc codant pour la GFP, notamment son mutant S65T ou S65C, dans lequel un ou plusieurs acides aminés sont substitués, insérés ou délétés afin de maintenir la phase de lecture et d'obtenir un polypeptide hybride.

5 L'invention a également pour objet des cellules contenant une séquence d'ADN codant pour un polypeptide, notamment un ligand, dans lequel un ou plusieurs acides aminés sont substitués, insérés ou délétés, fusionné avec l'ADNc codant pour la GFP, notamment son mutant S65T ou S65C, dans lequel un ou plusieurs acides aminés sont substitués, insérés ou délétés afin de maintenir la phase de lecture et capable d'exprimer la séquence d'ADN considérée.

10 L'invention a également pour objet la culture de cellules contenant l'hybride ligand-GFP dans des conditions permettant

- 1) l'expression du polypeptide hybride,
- 2) la détection de fluorescence de la cellule,
- 3) l'isolement de l'hybride ligand-GFP.

15 L'invention a également pour objet l'incubation de cellules exprimant l'hybride protéine-GFP, notamment un récepteur fusionné avec le mutant BFP (GFP Y145F Y66H) avec l'hybride ligand-GFP, notamment le mutant S65T ou S65C de la GFP et la mesure les changements de fluorescence soit d'émission du donneur (protéine-GFP), soit d'émission de l'accepteur (ligand-GFP), soit d'excitation de l'accepteur, révélant l'interaction protéine ligand et l'addition d'une molécule suspectée biologiquement active avec l'hybride ligand-GFP et la mesure des altérations du signal de transfert d'énergie par rapport à l'incubation des cellules exprimant l'hybride protéine-GFP et l'hybride ligand-GFP.

20 Ces différents aspects sont obtenus par l'utilisation d'une protéine fluorescente choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, notamment choisie parmi :

- la protéine fluorescente verte (GFP), ou
- des variants dérivés de la GFP, par addition, délétion ou substitution d'un ou plusieurs acides aminés, sous réserve que ces variants conservent la propriété de fluorescence,
- ou des fragments de la GFP, ou de fragments des susdits variants, sous réserve que ces fragments conservent la propriété de fluorescence,

pour la détection et la quantification d'interactions non covalentes entre une protéine cible marquée par la GFP ou l'un des variants ci-dessus définis ou l'un des fragments ci-dessus définis et l'un de ses ligands marqué par un marqueur constitué :

- 5 - soit par une molécule susceptible d'absorber la lumière émise par la protéine fluorescente,
- soit par une substance fluorescente,

cette détection et quantification ayant lieu par transfert d'énergie de fluorescence:

- 10 . entre la GFP ou l'un des variants définis ci-dessus, ou l'un des fragments définis ci-dessus et la susdite substance fluorescente, la substance fluorescente étant telle que soit elle est excitable à la longueur d'onde d'émission de la GFP ou de l'un des susdits variants, ou de l'un des susdits fragments, soit elle émet à la longueur d'excitation de la GFP, ou de l'un des susdits variants, ou de l'un des susdits fragments, ou
- 15 . entre la GFP ou l'un de ses variants définis ci-dessus, ou l'un des fragments définis ci-dessus, et la susdite molécule susceptible d'absorber la lumière émise par la protéine fluorescente.

20 L'invention a notamment pour objet l'utilisation d'une protéine fluorescente choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est supérieur à environ $14000\text{M}^{-1}\text{cm}^{-1}$ et le rendement quantique de fluorescence est supérieur à environ 0,38, notamment choisie parmi :

- 25 - la protéine fluorescente verte (GFP), ou
- des variants dérivés de la GFP, par addition, délétion ou substitution d'un ou plusieurs acides aminés, sous réserve que ces variants conservent la propriété de fluorescence,
- ou des fragments de la GFP, ou de fragments des susdits variants, sous réserve que ces fragments conservent la propriété de fluorescence,
- 30 pour la détection et la quantification d'interactions non covalentes entre une protéine cible marquée par voie génétique par la GFP ou l'un des variants ci-dessus définis ou l'un des fragments ci-dessus définis et l'un de ses ligands marqué par un marqueur constitué :

- 35 - soit par une molécule susceptible d'absorber la lumière émise par la protéine fluorescente,
- soit par une substance fluorescente,

cette détection et quantification ayant lieu par transfert d'énergie de fluorescence:

entre la GFP ou l'un des variants définis ci-dessus, ou l'un des fragments définis ci-dessus et la susdite substance fluorescente, la substance fluorescente étant telle que soit elle est excitable à la longueur d'onde d'émission de la GFP ou de l'un des susdits variants, ou de l'un des susdits fragments, soit elle émet à la longueur d'excitation de la GFP, ou de l'un des susdits variants, ou de l'un des susdits fragments, ou

entre la GFP ou l'un de ses variants définis ci-dessus, ou l'un des fragments définis ci-dessus, et la susdite molécule susceptible d'absorber la lumière émise par la protéine fluorescente.

La présente invention comporte les avantages suivants: La sensibilité de la détection de la cible et/ou du ligand obtenue par fusion avec la protéine fluorescente est telle que des concentrations de l'ordre du picomolaire sont aisément détectables. Par le fait que la fluorescence est introduite sur la cible (parfois aussi sur le ligand) par manipulation génétique, il résulte qu'il n'est pas nécessaire de disposer de la protéine que l'on souhaite rendre fluorescente, ce qui permet de détecter des protéines dont l'abondance naturelle empêche tout marquage chimique par manque de quantités manipulables (protéine non isolée ou non isolable) de celle-ci.

Le marquage par voie génétique offre en outre deux avantages supplémentaires, à savoir, quantité et homogénéité. Par opposition au marquage chimique qui peut survenir dans certains cas de manière partielle et inhomogène, la marquage par voie génétique est quantitatif et survient à une position prédéterminée et connue. Il en résulte que toutes les molécules sont marquées et qu'elles sont toutes marquées de manière identique.

L'invention consiste à détecter des interactions non covalentes entre une protéine et l'un de ses ligands, par transfert d'énergie de fluorescence, et à utiliser ce procédé dans le criblage de molécules biologiquement actives, notamment dans le domaine des récepteurs. La protéine est rendue fluorescente par fusion de son ADNc avec l'ADNc codant pour la GFP, le ligand est rendu fluorescent soit par greffage chimique d'un groupe fluorescent, soit par fusion de son gène avec celui de la GFP. L'interaction entre la protéine et son ligand entraîne des modifications du spectre de fluorescence de la GFP et/ou du ligand, qui peuvent être enregistrés en temps réel ou à l'équilibre.

L'interaction non covalente entre une protéine cible et son ligand correspond à la formation d'un complexe entre la protéine cible et le ligand

dans lequel la protéine cible et le ligand interagissent l'une avec l'autre, et sont retrouvés intacts après suppression de l'interaction

Une série de tests permettant de déterminer la nature non covalente, spécifique et saturable de l'interaction est notamment décrite dans les articles suivants : Levitski, A. 1980, in *Cellular receptors for hormones and neurotransmitters*, Eds Schulster, D. & Levitski, A., John Wiley & Sons Ltd.; Horovitz, H. and Levitski, A. 1987, *Proc. Natl. Acad. Sci.* 84:6654-6658; *Receptor Biochemistry and Methodology*, volume 3, Ventre, J.C. & Harrison, L.C. Eds, Alan R. Liss, INC, New York, 1987.

Dans le contexte présent, les termes suivants sont définis ainsi :

- protéine autofluorescente : protéine synthétique ou naturelle dans laquelle le chromophore se forme par une réaction autocatalytique entre des acides aminés de la protéine sans nécessiter l'addition d'un groupe prosthétique et dont les propriétés de fluorescence sont intrinsèques au monomère,

- "polypeptide hybride" : indique un polypeptide qui est une fusion d'au moins une partie de deux protéines, dans l'invention, à titre d'exemple au moins une partie de la GFP avec une partie d'une protéine cible ou une partie d'un ligand polypeptidique de la protéine cible,

- "compétiteur" : toute molécule se liant sur la protéine cible au même site que le ligand fluorescent,

- "substance biologiquement active" : toute substance susceptible d'interférer avec l'interaction d'un ligand fluorescent et de la protéine fluorescente spécifique en modifiant les paramètres cinétiques ou thermodynamiques de l'interaction.

On appelle ligand toute molécule qui interagit avec une autre molécule de manière non covalente et réversible.

L'expression "transfert d'énergie de fluorescence" correspond à un processus physique, dépendant de la distance, par lequel de l'énergie est transmise de manière non radiative d'un chromophore excité, le donneur, à un autre chromophore, l'accepteur, par interaction dipôle-dipôle (Förster 1951 in *Fluoreszenz organischer Verbindung*. Vandenhoek and Rupprecht, Göttingen; Wu and Brand 1994 *Anal. Biochem.* 218:1-13; Clegg 1995, *Current Opinion in Biotechnol.* 6:103-110). Le transfert d'énergie peut être observé soit par une diminution de l'amplitude de l'émission du donneur, soit par une augmentation de l'amplitude de l'excitation et de l'émission de l'accepteur.

Dans le cas de l'application du transfert d'énergie à des échantillons biologiques en interaction non covalentes, le signal de transfert ne peut pas persister si les conditions expérimentales ne permettent pas l'interaction entre

le ligand fluorescent et la protéine cible fluorescente. De même, si l'un des deux partenaires interagissant n'est pas fluorescent, les éventuelles variations de fluorescence observées pour l'autre partenaire ne pourront pas être attribuées à un processus de transfert d'énergie.

5 Les termes changement ou variation de fluorescence, définis dans le contexte du transfert d'énergie, se réfèrent à toutes modifications de 1) l'amplitude du signal de fluorescence du donneur, de 2) l'amplitude du spectre d'excitation ou 3) l'amplitude du signal d'émission du donneur. Les variations ou changements de fluorescence ne doivent pas être observés si l'un des deux partenaires n'est pas fluorescent ou si l'interaction entre les partenaires fluorescents est inhibée, par exemple par un excès d'un agent compétiteur.

10 De façon plus précise, la réaction de transfert d'énergie de fluorescence requiert deux groupes fluorescents, l'un appelé donneur et l'autre accepteur. Cette réaction se produit lorsque deux conditions sont réunies :

15 1) le spectre d'absorption de l'accepteur et le spectre d'émission du donneur doivent se recouvrir, au moins en partie ; le recouvrement se calcule à partir de données expérimentales et d'une équation donnant une valeur en cm^3M^{-1} (Lakey *et al.* 1991, *J. Mol. Biol.* 218:639-653) ;

20 2) le donneur et l'accepteur doivent être proches dans l'espace (de 10 à 100 angströms) afin que le transfert d'énergie puisse avoir lieu.

25 La première condition a pour conséquence le fait que l'excitation du donneur entraîne alors de manière concomitante une diminution de l'amplitude de l'émission du donneur et l'apparition d'un signal d'émission de l'accepteur. Ceci permet de détecter les interactions entre le donneur et l'accepteur et/ou de mesurer leur distance.

30 L'expression "proches dans l'espace" signifie que la distance entre le donneur et l'accepteur est inférieure ou égale à $2 R_0$, R_0 représentant le rayon de Forster (op.cit.) (Lakey, J.H. *et al.* 1991, *J. Mol. Biol.* 218:639-653).

35 Si l'accepteur n'est pas fluorescent, mais présente un spectre d'excitation recouvrant au moins en partie le spectre d'émission du donneur, le transfert d'énergie pourra être détecté sous la forme d'une réduction d'amplitude de l'émission du donneur.

L'invention concerne l'utilisation d'un ligand marqué par un marqueur constitué :

- 35 - soit par une molécule susceptible d'absorber la lumière émise par la protéine fluorescente,
- soit par une substance fluorescente,

pour la détection et la quantification d'interactions non covalentes entre une protéine cible et le susdit ligand, ladite protéine cible étant marquée par voie génétique par une protéine fluorescente choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est supérieur à environ $14000\text{M}^{-1}\text{cm}^{-1}$ et le rendement quantique de fluorescence est supérieur à environ 0,38, notamment choisie parmi :

- la protéine fluorescente verte (GFP), ou
- des variants dérivés de la GFP, par addition, délétion ou substitution d'un ou plusieurs acides aminés, sous réserve que ces variants conservent la propriété de fluorescence,
- ou des fragments de la GFP, ou de fragments des susdits variants, sous réserve que ces fragments conservent la propriété de fluorescence, cette détection et quantification ayant lieu par transfert d'énergie de fluorescence:
 - . entre la GFP ou l'un des variants définis ci-dessus, ou l'un des fragments définis ci-dessus et la susdite substance fluorescente, la substance fluorescente étant telle que soit elle est excitable à la longueur d'onde d'émission de la GFP ou de l'un des susdits variants, ou de l'un des susdits fragments, soit elle émet à la longueur d'excitation de la GFP, ou de l'un des susdits variants, ou de l'un des susdits fragments, ou
 - . entre la GFP ou l'un de ses variants définis ci-dessus, ou l'un des fragments définis ci-dessus, et la susdite molécule susceptible d'absorber la lumière émise par la protéine fluorescente.

20

Selon un mode de réalisation avantageux, la protéine fluorescente est choisie parmi :

- la protéine fluorescente verte (GFP) (Ward *et al.* 1980, Photochem. Photobiol. 31:611-615 ; Chalfie 1995, Photochem. Photobiol. 62:651-656), ou EGFP (Heim & Tsien, Current Biology, 1996, vol. n° 6, p178-182 ; Miyawaki *et al.*, Nature 1997, vol. 388, p882-887),
- la protéine fluorescente cyan (CFP ou ECFP) (Heim & Tsien, Current Biology, 1996, vol. n° 6, p178-182 ; Miyawaki *et al.*, Nature 1997, vol. 388, p882-887),
- la protéine fluorescente jaune (YFP ou EYFP) (Cormack *et al.* 1995, Gene 173:33-38 ; Heim, Cubitt et Tsien 1995, Nature ; Ehrig *et al.* 1995, FEBS Lett. 367:163-166) (Miyawaki *et al.*, Nature 1997, vol. 388, p882-887),

35

- GFPUV (Crameri *et al.* 1996, *Nature Biotechnol.* 14:315-319 ; Ehrig *et al.* 1995, *FEBS Lett.* 367:163-166),

ou leurs mutants dans lesquels les codons sont optimisés pour l'expression dans les cellules humaines, bactériennes ou végétales,

ou leurs mutants présentant des longueurs d'excitation ou d'émission plus élevées ou plus faibles que celles associées aux protéines définies ci-dessus, sous réserve que leur coefficient d'extinction moléculaire soit supérieur à environ $14000M^{-1}cm^{-1}$ et leur rendement quantique de fluorescence soit supérieur à environ 0,38.

L'expression "codons optimisés" indique la substitution de codons de la protéine sauvage par leurs homologues préférés de l'organisme hôte, sans changement de code donc sans changement de séquence protéique.

La GFP sauvage (WT) de longueur d'onde, d'excitation et d'émission 395/470-509 est décrite dans ref: Ward *et al.* 1980 *Photochem. Photobiol.* 31:611-615, Chalfie 1995, *Photochem. Photobiol.* 62:651-656.

La GFP UV présentant les mutations suivantes : F99S, M153T, V163A de longueur d'onde, d'excitation et d'émission respectivement de 395-510 est décrite dans Crameri *et al.* 1996 *Nature Biotechnol.* 14:315-319. ou avec la mutation T203I et la longueur d'onde respectivement d'excitation et d'émission 400-512 est décrite dans Ehrig *et al.* 1995 *FEBS Lett.* 367:163-166.

La EGFP présente les mutations suivantes :

F64L S65T H231L

La EYFP présente les mutations suivantes :

S65G V68L S72A T203Y

La ECFP présente les mutations suivantes :

F64L	S65T	Y66W	N146I
M153T	V163A	N212K	

Les différents mutants de GFP peuvent en outre être optimisés (par l'introduction de mutations silencieuse optimisant l'usage de codons spécifiques à chaque espèce) pour l'expression dans ces cellules :

- humaines, ref Haas *et al.* 1996 *Curr. Biol.* 6:315-323; Yan *et al.* 1996, *Nucleic Ac. Res.* 24:4592-4593; Zolotukhin *et al.* 1996, *J. Virol.* 70:4646-4654
- bactériennes : Crameri *et al.* 1996 *Nature Biotechnol.* 14:315-319, Cormack *et al.* 1996, *Gene*, 173:33-38 pour *Escherichia coli*,
- végétales: Reichel *et al.* 1996, *Proc. Natl. Acad. Sci.* 93:5888-5893.

. GFP :

Le terme GFP indique une protéine codée par la séquence nucléotidique donnée en figure 1, et qui une fois exprimée dans des cellules émet une fluorescence. Les GFPs présentant des substitutions, additions ou délétions d'acides aminés influençant soit les propriétés de fluorescence, soit le taux d'expression de la GFP sont appelés mutants de GFP.

On donne ci-après les principales caractéristiques des protéines fluorescentes avantageusement utilisées dans le procédé de l'invention :

Protéine	λ -excitation maximale	λ -émission	coefficient d'extinction	rendement quantique
EYFP	514	527	36500	0,63
ECFP	432	480	18000	0,67
GFPUV	395	509	21000	0,77
EGFP	489	511	39000	0,66

10

La protéine autofluorescente BFP est de préférence exclue car elle ne répond pas aux conditions définies pour les protéines autofluorescentes de cnidaire, à savoir coefficient d'extinction moléculaire supérieur à $14000\text{ M}^{-1}\text{cm}^{-1}$ et rendement quantique de fluorescence supérieur à 0,38.

15

L'invention concerne également l'utilisation d'une protéine fluorescente (n° 1) telle que définie ci-dessus, dans laquelle le ligand est marqué

* soit par une substance fluorescente, le marquage étant :

- soit effectué par voie chimique, la substance fluorescente étant alors un composé chimique,

20

- soit effectué par voie recombinante, la substance fluorescente étant alors un peptide ou une protéine fluorescente (n° 2), et pouvant être notamment choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est supérieur à environ $14000\text{ M}^{-1}\text{cm}^{-1}$ et le rendement quantique de fluorescence est supérieur à environ 0,38, notamment choisie parmi :

25

- la protéine fluorescente verte (GFP), ou
 - des variants dérivés de la GFP, par addition, délétion ou substitution d'un ou plusieurs acides aminés, sous réserve que ces variants conservent la propriété de fluorescence,

30

- ou des fragments de la GFP, ou de fragments des susdits variants, sous réserve que ces fragments conservent la propriété de fluorescence,

* soit par une substance non fluorescente appartenant au groupe des acides violets [Acid Violet 5, CAS 10130-48-0 ; Acid Violet 7, CAS 4321-69-1 ; Acid Violet 17, CAS 4129-84-4], acides rouges [Acid Red 1, CAS 3734-67-6 ; Acid Red 8, CAS 4787-93-3 ; Acid Red 37, CAS 6360-07-2 ; Acid Red 40, CAS 12167-45-2 ; Acid Red 106, CAS 6844-74-2 ; Acid Red 114, CAS 6459-94-5], les alizarines, l'aluminon, l'azocarmine B [CAS 25360-72-9], la fuschine basique [Basic Red 9, CAS 569-61-9], le Bordeaux R [Acid Red 17, CAS 5858-33-3], la Carmine [CAS 1390-65-4].

"CAS" correspond à Chemical Abstracts.

10 Marquage :

Par marquage d'une protéine cible ou d'un ligand, on entend:

- pour la protéine cible, la fusion de son gène ou ADNc, ou partie du gène ou de l'ADNc, avec le gène ou ADNc, ou partie du gène ou ADNc, de la GFP;
- 15 - pour le ligand, il peut s'agir d'un couplage chimique entre le ligand et un groupe fluorescent, ou bien de la fusion de son gène ou ADNc, ou partie du gène ou de l'ADNc, avec le gène ou ADNc, ou partie du gène ou ADNc, de la GFP.

20 L'invention concerne l'utilisation d'une protéine fluorescente selon l'invention dans laquelle la protéine cible et le ligand sont marqués par voie génétique, la protéine fluorescente et la substance fluorescente étant respectivement choisies parmi les couples de composés suivants :

25 GFPuv - EYFP
EYFP - GFPuv
ECFP - EYFP
EYFP - ECFP
ECFP - EGFP
EGFP - ECFP
EGFP - EYFP
30 EYFP - EGFP

et notamment dans laquelle la protéine cible est marquée par la protéine EYFP ou EGFP et le ligand est marqué par la protéine ECFP, ou la protéine cible est marquée par la protéine ECFP et le ligand est marqué par la protéine EYFP ou EGFP.

35 L'invention a également pour objet l'utilisation d'une protéine fluorescente choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est

supérieur à environ $14000\text{M}^{-1}\text{cm}^{-1}$ et le rendement quantique de fluorescence est supérieur à environ 0,38, notamment choisie parmi :

- la protéine fluorescente verte (GFP), ou
- des variants dérivés de la GFP, par addition, délétion ou substitution d'un ou plusieurs acides aminés, sous réserve que ces variants conservent la propriété de fluorescence,
- ou des fragments de la GFP, ou de fragments des susdits variants, sous réserve que ces fragments conservent la propriété de fluorescence,

pour la détection et la quantification d'interactions non covalentes entre une protéine cible marquée par la GFP ou l'un des variants ci-dessus définis ou l'un des fragments ci-dessus définis et l'un de ses ligands marqué par une substance fluorescente, cette détection et quantification ayant lieu par transfert d'énergie de fluorescence entre la GFP ou l'un des variants définis ci-dessus, ou l'un des fragments définis ci-dessus et la susdite substance fluorescente, la substance fluorescente étant telle que soit elle est excitable à la longueur d'onde d'émission de la GFP ou de l'un des susdits variants, ou de l'un des susdits fragments, soit elle émet à la longueur d'onde d'excitation de la GFP, ou de l'un des susdits variants, ou de l'un des susdits fragments.

Selon un mode de réalisation avantageux de l'invention, la protéine fluorescente est EGFP et dans laquelle :

- soit la EGFP est donneur d'énergie de fluorescence et le marqueur absorbant la lumière émise par la EGFP est une substance fluorescente ou non, et le marqueur étant choisi parmi des substances, dont le spectre d'excitation chevauche le spectre d'émission de la EGFP, et notamment dans le cas où le marqueur est une substance fluorescente, il est choisi parmi : le 4,4-difluoro-4-bora-3a,4a-diaza-s-indacene (Bodipy), l'éosine, l'érythrosine, la tétraméthylrhodamine, la sulforhodamine 101 commercialisée par Molecular probe sous la dénomination Texas Red, et leurs dérivés permettant d'une part le greffage et, d'autre part, dont le spectre d'excitation recouvre le spectre d'émission de EGFP,

et dans le cas où le marqueur n'est pas une substance fluorescente, il est choisi parmi le groupe des acides violets [Acid Violet 5, CAS 10130-48-0 ; Acid Violet 7, CAS 4321-69-1 ; Acid Violet 17, CAS 4129-84-4], des acides rouges [Acid Red 1, CAS 3734-67-6 ; Acid Red 8, CAS 4787-93-3 ; Acid Red 37, CAS 6360-07-2 ; Acid Red 40, CAS 12167-45-2 ; Acid Red 106, CAS 6844-74-2 ; Acid Red 114, CAS 6459-94-5], les alizarines, l'aluminon, l'azocarmine B [CAS 25360-72-9], la fuschine

basique [Basic Red 9, CAS 569-61-9], le Bordeaux R [Acid Red 17, CAS 5858-33-3], la Carmine [CAS 1390-65-4],

- soit la EGFP est accepteur d'énergie de fluorescence et la substance fluorescente est donneur d'énergie de fluorescence et est choisie parmi des substances, dont le spectre d'émission chevauche le spectre d'excitation de la EGFP, et notamment parmi : les coumarines, la fluorescamine, le 6-(N-méthylanilino)naphtalène, (mansyl) et leurs dérivés permettent d'une part le greffage et, d'autre part, dont le spectre d'émission recouvre le spectre d'excitation de EGFP,

10 - ou soit la protéine fluorescente est ECFP et est donneur d'énergie de fluorescence et la substance fluorescente est accepteur d'énergie et est choisi parmi la fluoresceïne et le 7-nitrobenz-2-oxa-1,3-diazole,

15 - ou soit la protéine fluorescente est ECFP et est accepteur d'énergie de fluorescence et la substance fluorescente est donneur d'énergie et est choisie parmi le pyrène ou la coumarine ou leurs dérivés permettant d'une part le greffage, et, d'autre part, dont le spectre d'émission chevauche le spectre d'excitation de la ECFP.

S'agissant de la protéine cible, elle peut être choisie parmi :

20 - les récepteurs membranaires couplés à la protéine G, notamment dans Supplement Trends in Pharmacological Sciences, 1997 (*Receptor and ion Channel Nomenclature*),

25 - les récepteurs des facteurs de croissance, notamment ceux qui sont structurellement reliés au récepteur de l'insuline (Yarden, Y. and Ullrich, A. 1988, Biochemistry 27:3113-3119) ou au récepteur de l'interféron γ (Brisco, J. et al. 1996, Phylos. Trans. R. Soc. Lond. B. Biol. Sci. 351:167-171 ; Ihle, J.N. 1995, Nature 377:591-594),

30 - les récepteurs canaux, notamment dans Supplement Trends in Pharmacological Sciences, 1997 (*Receptor and ion Channel Nomenclature*),

35 - les récepteurs nucléaires intracellulaires, notamment ceux qui sont structurellement reliés au récepteur des stéroïdes (Mangelsdorf et al. 1995, Cell, 83:835-839 ; Wurtz, J.L. et al. 1996, Nature Struct. Biol. 3:206).

Selon un mode de réalisation avantageux, la protéine cible est choisie parmi les récepteurs membranaires couplés à la protéine G.

Dans ce qui précède et ce qui suit, on désigne par "récepteur" toute molécule de nature protéique susceptible d'entrer en interaction non covalente avec un agent pharmacologique. Préférentiellement dans l'invention un récepteur de neurotransmetteur, d'hormones, de facteur de croissance etc.,

capable de produire, après interaction avec un ligand pharmacologique, une réponse de transduction de signal mesurable *in vivo* et/ou *in vitro*.

Par réponse de transduction de signal, on désigne toute réponse, ou inhibition de réponse, mesurable *in vivo* et/ou *in vitro*, résultant de l'interaction d'un récepteur avec ses agent pharmacologiques spécifiques et conduisant à des activations, ou inhibitions, du métabolisme cellulaire par effet sur des messagers secondaires, des enzymes, ou des courants ioniques.

S'agissant de la réponse de transduction de signal pour les récepteurs couplés aux protéines G, le test général consiste à déterminer l'activation de la protéine G par mesure de la fixation de GTP (Befort *et al.* 1996 Neurochem. Res. 11:1301-1307). D'autres mesures plus spécifiques font par exemple intervenir des déterminations de concentrations intracellulaires d'AMPc, d'inositol phosphates, de calcium, des mesures d'activation de la transcription de gènes ou d'activité oncogénique, suivant le type de couplage spécifique du récepteur considéré.

Pour les récepteurs-canaux, les mesures les plus directes sont des déterminations de courants ioniques (Hille, B. 1992 in Ion channels of excitable membranes, Sinauer Associates, Sunderland, Massachussets). D'autres mesures peuvent, par exemple, faire intervenir des détermination des transcription de gènes ou des activations d'enzymes.

Pour les récepteurs de facteurs de croissance, les tests généraux sont ceux de la prolifération, de différentiation ou de la survie cellulaire, fréquemment aussi des tests de phosphorylations de substrats spécifiques (Honneger *et al.* 1988, EMBO J. 7:3053-3060) de chaque récepteur et repérage par des anticorps spécifiques de phosphoaminoacides.

Pour les récepteurs nucléaires, les tests de transduction de signal sont ceux de la transcription de gènes dans lesquels des gènes rapporteurs "chromogéniques" sont placés sous le contrôle de promoteurs spécifiques des voies de transduction du récepteur étudié.

A titre d'exemples de récepteurs membranaires couplés aux protéines G, on peut citer les récepteurs de purines et nucléotides, des amines biogéniques des peptides et protéines, des eicosanoides, des lipides et dérivés, des acides aminés excitateurs et des ions, des molécules olfactives ainsi que les récepteurs orphelins (ci-après une liste assez exhaustive).

A titre d'exemple de récepteurs de facteurs de croissance, on peut citer les cytokines, le facteur de croissance épidermique, l'insuline, le facteur de croissance dérivé des plaquettes, le facteur de croissance de transformation.

A titre de récepteurs canaux, on peut citer notamment les récepteurs de l'ATP, de la sérotonine, du GABA, de la glycine, de l'acétylcholine, du glutamate.

5 A titre d'exemple de récepteurs nucléaires, on peut citer notamment les récepteurs des hormones thyroidiennes, des oestrogènes, des glucocorticoïdes, des rétinoïdes.

A titre de ligands des récepteurs couplés à la protéine G on peut citer :

10 ● Purines et Nucléotides

- . Adénosine
- . cAMP
- . ATP
- . UTP
- . ADP

15 ● Amines Biogéniques (et ligands naturels reliés)

- . 5-hydroxytryptamine
- . Acétylcholine
- . Dopamine
- . Adrénaline
- . Histamine
- . Mélatonine
- . Noradrénaline
- . Tyramine/Octopamine
- . autres composés reliés

20 ● Peptides

- . Hormone adrénocorticotrophique (ACTH)
- . Hormone stimulatrice de mélanocyte (MSH)
- . Mélanocortines
- . Neurotensine (NT)
- . Bombésine et peptides voisins
- . Endothélines
- . Cholecystokinine
- . Gastrine
- . Neurokinine B (NKB)
- . Récepteur des tachykinines
- . Substance K (NKA)

- . Substance P (SP)
- . Neuropeptide Y (NPY)
- . Facteur de libération de la thyrotropine
- . Nociceptine
- 5 . Bradykinine
- . Angiotensine II
- . Beta-endorphine
- . C5a anaphalatoxine
- . Calcitonine
- 10 . Chemokines (également appelés intercrines)
- . Facteur de libération corticotrophique (CRF)
- . Dynorphine
- . Endorphine
- . Peptides formylés
- 15 . Follitropine (FSH)
- . Phéromones de maturation fongique
- . Galanine
- . Récepteur du polypeptide inhibiteur gastrique (GIP)
- . Peptides analogues du glucagon (GLPs)
- 20 . Glucagon
- . Hormone de libération de gonadotropine (GmRH)
- . Hormone de libération de l'hormone de croissance (GHRM)
- . Hormone diurétique d'insecte
- . Interleukine
- 25 . Leutropine (LH/HCG)
- . MET -enképhaline
- . Peptides opioides
- . Oxytocine
- . Hormone parathyroïde (PTH) et (PTHrP)
- 30 . Peptides activant l'adényl cyclase pituitaire (PACAP)
- . Secretine
- . Somatostatine
- . Thrombine
- . Thyrotropine (TSH)
- 35 . Peptide intestinal vasoactif (VIP)
- . Vasopressine
- . Vasotocine

- Eicosanoides
 - . IP - Prostacyclines
 - . PG - Prostaglandines
 - . TX - Thromboxanes
- 5 ● Composés à base de rétinal
 - . 11-cis retinal de vertébré
 - . 11-cis retinal d'invertébré
- 10 ● Lipides et composés à bases de lipides
 - . Cannabinoides
 - . Anandamide
 - . Acide lysophosphatidique
 - . Facteur d'activation des plaquettes
- 15 ● Leukotriènes
- Amino acides excitateurs et ions
 - . Ion Calcium
 - . Glutamate
- 20 ● Récepteurs orphelins
 - . Récepteurs olfactifs
- Le ligand peut être soit un agoniste, soit un antagoniste.
Par "agoniste", on entend toute molécule mimant l'effet du ligand endogène naturel, par exemple le neurotransmetteur, le facteur de croissance, l'hormone.
- 25 Par "antagoniste", on entend toute molécule inhibant l'effet de l'agoniste en se liant sur la même protéine cible que ce dernier.
- L'invention concerne un procédé de détection et de quantification d'interactions non covalentes entre une protéine cible, notamment un récepteur, et l'un de ses ligands, caractérisé en ce que :
30 - on prépare des cellules soit des fragments de cellules contenant une séquence d'ADN comprenant le gène codant pour une protéine fluorescente fusionnée avec le gène de la protéine cible, la fusion entre le gène de la protéine fluorescente et le gène de la susdite protéine cible étant telle que les propriétés de la protéine cible, notamment du récepteur, ne sont pas modifiées par la présence de la protéine fluorescente, à savoir :
- 35

- * l'interaction entre la protéine cible, notamment le récepteur, et le ligand n'est pas modifiée,
- * la fonction de transduction de la réponse n'est pas modifiée, la protéine fluorescente étant choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est supérieur à environ $14000M^{-1}cm^{-1}$ et le rendement quantique de fluorescence est supérieur à environ 0,38, notamment choisie parmi :

10 - la protéine fluorescente verte (GFP), ou
- des variants dérivés de la GFP, par addition, délétion ou substitution d'un ou plusieurs acides aminés, sous réserve que ces variants conservent la propriété de fluorescence,

15 - ou des fragments de la GFP, ou de fragments des susdits variants, sous réserve que ces fragments conservent la propriété de fluorescence,

- on met en présence les susdites cellules ou les susdits fragments de cellules avec un ligand de la susdite protéine cible, notamment du susdit récepteur, marqué par un marqueur constitué :

- soit par une molécule susceptible d'absorber la lumière émise par la protéine fluorescente,

- soit par une substance fluorescente,

et soit la protéine fluorescente étant donneur d'énergie de fluorescence et le marqueur étant accepteur d'énergie de fluorescence, ou soit la protéine fluorescente étant accepteur d'énergie de fluorescence et le marqueur étant une substance fluorescente donneur d'énergie de fluorescence, et

- on irradie à une longueur d'onde permettant soit d'exciter la protéine fluorescente, soit d'exciter la substance fluorescente,

- les susdites étapes de mise en présence et d'irradiation pouvant être effectuées soit simultanément, soit l'une après l'autre, ou

30 - on met en présence les susdites cellules ou les susdits fragments de cellules avec un ligand de la susdite protéine, notamment du susdit récepteur, marqué par un marqueur, les cellules ou le ligand ayant été irradiés préalablement à leur mise en présence,

- on détecte soit une diminution d'amplitude de l'émission du donneur et/ou un signal d'émission caractéristique de l'émission de l'accepteur.

. Cellules, fragments et protéine purifiée :

Par cellule, on entend toute cellule eucaryote ou procaryote, animale ou végétale dans laquelle on introduit le gène codant pour une protéine de fusion entre un récepteur (déscrits) et la GFP ou ses mutants ou fragments.

5 Par fragment de cellule, on entend toute fraction, membranaire ou non membranaire, obtenue à partir des cellules et contenant la protéine de fusion récepteur-GFP.

Par protéine purifiée, on entend protéine de fusion partiellement ou totalement purifiée.

10 Selon un mode de réalisation avantageux, le procédé utilise l'ADNc du variant EGFP de la protéine de fluorescence verte fusionnée à un récepteur en association avec un ligand dudit récepteur marqué par le Bodipy 530-550. Le système d'expression préféré est la cellule mammifère ou la levure.

15 L'invention a également pour objet un procédé de détection et de quantification d'interactions non covalentes entre une protéine cible, notamment un récepteur, et l'un de ses ligands, caractérisé en ce que :

20 - on prépare une protéine fluorescente fusionnée avec une protéine cible dont on veut déterminer l'interaction protéine-ligand, la fusion entre la protéine fluorescente et la susdite protéine cible étant telle que les propriétés de la protéine, notamment du récepteur, ne sont pas modifiées par la présence de la protéine fluorescente, à savoir :

* l'interaction entre la protéine cible, notamment le récepteur, et le ligand n'est pas modifiée,

25 * la fonction de transduction de la réponse n'est pas modifiée, la protéine fluorescente étant choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est supérieur à environ $14000\text{M}^{-1}\text{cm}^{-1}$ et le rendement quantique de fluorescence est supérieur à environ 0,38, notamment choisie parmi :

30 - la protéine fluorescente verte (GFP), ou
- des variants dérivés de la GFP, par addition, délétion ou substitution d'un ou plusieurs acides aminés, sous réserve que ces variants conservent la propriété de fluorescence,

- ou des fragments de la GFP, ou de fragments des susdits variants, sous réserve que ces fragments conservent la propriété de fluorescence,

35 - on met en présence la susdite protéine fluorescente fusionnée avec la protéine cible avec un ligand de la susdite protéine, notamment du susdit récepteur, ce ligand étant marqué par un marqueur constitué :

- soit par une molécule susceptible d'absorber la lumière émise par la protéine fluorescente,

- soit par une substance fluorescente,

et soit la protéine fluorescente étant donneur d'énergie de fluorescence et le marqueur étant accepteur d'énergie de fluorescence, ou soit la protéine fluorescente étant accepteur d'énergie de fluorescence et le marqueur est une substance fluorescente donneur d'énergie de fluorescence, et

- on irradie à une longueur d'onde permettant soit d'exciter la protéine fluorescente, soit d'exciter la substance fluorescente,

10 - les susdites étapes de mise en présence et d'irradiation pouvant être effectuées soit simultanément, soit l'une après l'autre, ou

- on met en présence la susdite protéine fluorescente fusionnée avec la protéine cible avec un ligand de la susdite protéine, notamment du susdit récepteur, ce ligand étant marqué par un marqueur constitué :

15 - soit par une molécule susceptible d'absorber la lumière émise par la protéine fluorescente,

- soit par une substance fluorescente,

la protéine fluorescente fusionnée avec la protéine cible ou le ligand ayant été irradiés préalablement à leur mise en présence,

20 - on détecte soit une diminution d'amplitude de l'émission du donneur et/ou un signal d'émission caractéristique d'émission de l'accepteur.

La protéine cible et notamment le récepteur modifié par fusion avec la GFP devra, outre exhiber la fluorescence de la GFP, conserver des propriétés pharmacologiques compatibles avec sa définition de récepteur spécifique des composés le caractérisant. En particulier, il devra pouvoir lier les même ligands que le récepteur sauvage.

25 Dans le procédé de l'invention, la mise en présence de la protéine fluorescente fusionnée avec la protéine cible et du ligand, ou de cellules ou fragments de cellules (contenant la susdite protéine fluorescente fusionnée avec la protéine cible) et du ligand conduit à un mélange.

30 Le mélange peut être effectué dans n'importe quel ordre d'addition. L'ordre préférentiel sera l'addition de ligand fluorescent à une solution de cellules ou de protéine, si la GFP est donneur, l'inverse si la GFP est accepteur.

Le mélange peut être effectué dans un appareil de mélange rapide combiné 35 à un système de détection de fluorescence pour avoir accès à des mesures en temps réel de l'interaction récepteur ligand et de ses altérations par des molécules biologiquement actives.

. Introduction d'ADN dans les cellules :

5 Les cellules animales, végétales, d'insectes, les levures, les bactéries ou les champignons constituant le système d'expression choisi pourront soit être transfectées, transformées, électroporées, ou infectées, de manière stable ou transitoire selon les protocoles décrits dans les manuels de laboratoire, par exemple Current Protocols in Molecular Biology, eds Ausabel *et al.*, John Wiley and sons.

L'invention a également pour objet un procédé dans lequel on introduit une étape supplémentaire :

10 - avant ou après, ou simultanément à l'étape de mise en présence de la protéine fluorescente fusionnée avec la protéine cible et d'un ligand de la susdite protéine, ce ligand étant marqué par un marqueur, ou

15 - avant, ou après, ou simultanément à l'étape de mise en présence de cellules ou de fragments de cellules et d'un ligand de la susdite protéine, marqué par un marqueur,

cette étape supplémentaire consistant :

- soit à mettre la susdite protéine fluorescente fusionnée avec la protéine cible en présence du susdit ligand non marqué et simultanément du susdit ligand marqué,

20 - soit à mettre en présence les susdites cellules ou les susdits fragments de cellules en présence simultanée du susdit ligand non marqué et du susdit ligand marqué,

25 - on détecte soit une diminution d'amplitude de l'émission du donneur et/ou un signal d'émission caractéristique de l'émission de l'accepteur, respectivement dans le cas de l'utilisation du ligand marqué et dans le cas de l'utilisation simultanée du ligand marqué et du ligand non marqué,

- et on compare soit les diminutions d'amplitude de l'émission du donneur respectivement obtenues et/ou les signaux d'émission caractéristiques de l'émission de l'accepteur respectivement obtenus.

30 Dans ce cas, il y a compétition entre le ligand marqué et le ligand non marqué, vis à vis de la protéine cible. L'intérêt d'un tel procédé est de pouvoir être adapté sur des automates pour identifier de nouvelles molécules capables d'interagir avec la protéine cible et se comportant soit comme agoniste soit comme antagoniste vis-à-vis de la réponse biologique. L'analyse, à différentes concentrations de la molécule non marquée, de l'inhibition de l'interaction entre la protéine cible et son ligand fluorescent permet par ailleurs de déterminer les paramètres d'affinité de la molécule non marquée pour la protéine cible.

L'invention a également pour objet un procédé, dans lequel la protéine fluorescente est EGFP et dans lequel :

- soit la EGFP est donneur d'énergie de fluorescence et le marqueur est accepteur d'énergie de fluorescence et est choisie parmi des substances, dont le spectre d'excitation chevauche le spectre d'émission de la EGFP, et notamment dans le cas où le marqueur est une substance fluorescente, il est choisi parmi : le 4,4-difluoro-4-bora-3a,4a-diaza-s-indacene (Bodipy), l'éosine, l'érythrosine, la tétraméthylrhodamine, la sulforhodamine 101 commercialisée par Molecular probe sous la dénomination Texas Red, et leurs dérivés permettent d'une part le greffage et, d'autre part, dont le spectre d'excitation recouvre le spectre d'émission de EGFP,

et dans le cas où le marqueur n'est pas une substance fluorescente, il est choisi parmi le groupe des acides violets [Acid Violet 5, CAS 10130-48-0 ; Acid Violet 7, CAS 4321-69-1 ; Acid Violet 17, CAS 4129-84-4], acides rouges [Acid Red 1, CAS 3734-67-6 ; Acid Red 8, CAS 4787-93-3 ; Acid Red 37, CAS 6360-07-2 ; Acid Red 40, CAS 12167-45-2 ; Acid Red 106, CAS 6844-74-2 ; Acid Red 114, CAS 6459-94-5], les alizarines, l'aluminon, l'azocarmine B [CAS 25360-72-9], la fuschine basique [Basic Red 9, CAS 569-61-9], le Bordeaux R [Acid Red 17, CAS 5858-33-3], la Carmine [CAS 1390-65-4],

- soit la EGFP est accepteur d'énergie de fluorescence et la substance fluorescente est donneur d'énergie de fluorescence et est choisie parmi des substances, dont le spectre d'émission chevauche le spectre d'excitation de la EGFP, et notamment parmi : les coumarines, la fluorescamine, le 6-(N-méthylanilino)naphtalène, (mansyl) et leurs dérivés permettent d'une part le greffage et, d'autre part, dont le spectre d'émission recouvre le spectre d'excitation de EGFP.

L'invention a également pour objet un procédé dans lequel la protéine dont on veut déterminer l'interaction protéine-ligand est choisie parmi :

- les récepteurs membranaires couplés à la protéine G, notamment dans Supplement Trends in Pharmacological Sciences, 1997 (*Receptor and ion Channel Nomenclature*),

- les récepteurs des facteurs de croissance, notamment ceux qui sont structurellement reliés au récepteur de l'insuline (Yarden, Y. and Ullrich, A. 1988, Biochemistry 27:3113-3119) ou au récepteur de l'interféron γ (Brisco, J. et al. 1996, Phylos. Trans. R. Soc. Lond. B. Biol. Sci. 351:167-171 ; Ihle, J.N. 1995, Nature 377:591-594),

- les récepteurs canaux, notamment dans *Supplement Trends in Pharmacological Sciences*, 1997 (*Receptor and ion Channel Nomenclature*),

- les récepteurs nucléaires intracellulaires, notamment ceux qui sont structurellement reliés au récepteur des stéroïdes (Mangelsdorf *et al.* 1995, *Cell*, 83:835-839 ; Wurtz, J.L. *et al.* 1996, *Nature Struct. Biol.* 3:206).

5

L'invention a également pour objet un procédé dans lequel la protéine fluorescente est la EGFP et la substance marquée est le Bodipy et dans lequel on détecte soit la diminution d'amplitude d'émission de la EGFP, soit le signal d'émission du Bodipy résultant du transfert d'énergie, la longueur d'onde d'irradiation correspondant à la longueur d'onde d'excitation de la EGFP.

10

L'invention a également pour objet un procédé dans lequel la protéine fluorescente est la EGFP et la substance marquée est une coumarine et dans lequel on détecte soit la diminution d'amplitude de la coumarine, soit le signal d'émission de la EGFP résultant du transfert d'énergie, la longueur d'onde d'irradiation correspondant à la longueur d'onde d'excitation de la coumarine.

15

L'invention concerne également un procédé dans lequel la protéine fluorescente est fusionnée du côté N-terminal et la protéine cible, notamment le récepteur, est fusionné du côté C-terminal.

20

L'invention concerne également un procédé dans lequel la protéine fluorescente est fusionnée du côté C-terminal et la protéine cible, notamment le récepteur, est fusionné du côté N-terminal.

25

L'invention a également pour objet un procédé dans lequel la protéine fluorescente est insérée dans la protéine cible à un endroit ne correspondant pas à un site de liaison protéine cible-ligand, notamment dans le cas des récepteurs couplés à la protéine G, cette insertion ayant lieu dans la première ou la troisième boucle intracellulaire du récepteur, sous réserve que l'insertion ne détruise ni les propriétés du récepteur, ni la fluorescence de la protéine fluorescente.

30

L'invention concerne également un procédé dans lequel les cellules sont des cellules de mammifères, notamment les cellules HEK 293 adhérentes ou en suspension, cellules CHO, cellules COS, lignées lymphocytaires, fibroblastes, etc., ou des cellules de levure, notamment *pichia* telle que *pichia pastoris*, *saccharomyces* telle que *saccharomyces cerevisiae*, *saccharomyces kluyveri*, *Hansenula* telle que *hansenula polymorpha*, ou des cellules d'insectes infectées par un virus tel que *baculovirus*, notamment cellules TNI ou sf9, ou des champignons, notamment les souches de *Aspergillus* (*A. oryzae*, *A. nidulans*, *A. niger*), *Neurospora*, *Fusarium*, *Trichoderma*.

35

5 L'utilisation de *Aspergillus* pour l'expression de protéines est décrite dans EP 0,272,277 ou EP 0,230,023 ou EP 0,184,438, ou des cellules de plantes, notamment *Arabidopsis* (*A. thaliana*), ou des protoplastes de citrange (*Citrus sinensis*) est décrite dans Haseloff, J. and Amos, B. 1995, Trends in Genetics 11:328-329.

S'agissant de la lignée cellulaire utilisée, elle dépend du récepteur; il est en effet souhaitable de choisir une lignée n'exprimant pas déjà de manière naturelle le récepteur choisi.

10 L'invention a également pour objet un procédé dans lequel un signal est détectable, dans un appareil de fluorimétrie conventionnel ou dans un appareil de mélange rapide équipé d'un système de détection de fluorescence, après mélange du donneur et de l'accepteur et peut être aboli par l'addition d'une substance non fluorescente de même spécificité pharmacologique, et notamment dans lequel le rapport signal/bruit est supérieur à environ 2.

15 Pour fixer les idées et à titre d'exemple, le rapport signal/bruit est d'environ 100 avec le couple BODIPY-EGFP sur l'équipement utilisé (décrit dans les exemples) et avec le récepteur de la substance qu'on désignera par NK2R marquée par la EGFP (NK2R-EGFP) et son ligand, la substance K (NKA) marquée par le BODIPY (NKA-BO).

20 Concernant le signal et le rapport signal sur bruit, il peut être défini comme suit :

25 Si le mélange du donneur et de l'accepteur est accompagné d'un changement de fluorescence, soit du donneur, soit de l'accepteur, ce changement doit être inhibé et renversé par une substance non fluorescente (de même spécificité pharmacologique) et qui de plus définit la spécificité du phénomène pharmacologique observé. Pour certaines applications (réponse de type oui-non) un faible rapport signal sur bruit peut suffire.

30 S'agissant de la protéine cible, il est possible de faire des mesures avec des cellules entières (contenant l'ADN codant pour la protéine cible), des fragments de cellules (membranes pour des récepteurs membranaires) des échantillons de protéines cibles solubilisés (récepteurs membranaires) ou purifiés (cf. exemples).

35 L'invention a également pour objet l'utilisation d'un protéine fluorescente choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est supérieur à environ $14000M^{-1}cm^{-1}$ et le rendement quantique de fluorescence est supérieur à environ 0,38, notamment choisie parmi :

- la protéine fluorescente verte (GFP), ou

- des variants dérivés de la GFP, par addition, délétion ou substitution d'un ou plusieurs acides aminés, sous réserve que ces variants conservent la propriété de fluorescence,

5 - ou des fragments de la GFP, ou de fragments des susdits variants, sous réserve que ces fragments conservent la propriété de fluorescence,

pour la détection et la quantification d'interactions non covalentes entre une protéine cible constituée par un récepteur couplé aux protéines G et une protéine G, en vue d'identifier les molécules biologiquement actives vis-à-vis du récepteur, susceptibles de former une interaction non covalente réversible avec ledit récepteur, ledit récepteur étant marqué par voie génétique par la protéine fluorescente et la protéine G étant marquée par un marqueur constitué :

10 - soit par une molécule susceptible d'absorber la lumière émise par la protéine fluorescente,

15 - soit par une substance fluorescente pouvant être notamment choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est supérieur à environ $14000M^{-1}cm^{-1}$ et le rendement quantique de fluorescence est supérieur à environ 0,38, notamment choisie parmi :

20 - la protéine fluorescente verte (GFP), ou

25 - des variants dérivés de la GFP, par addition , délétion ou substitution d'un ou plusieurs acides aminés, sous réserve que ces variants conservent la propriété de fluorescence,

30 - ou des fragments de la GFP, ou de fragments des susdits variants, sous réserve que ces fragments conservent la propriété de fluorescence,

35 cette détection et quantification ayant lieu par transfert d'énergie de fluorescence entre le récepteur marqué par la GFP ou l'un de ses variants définis ci-dessus, ou l'un des fragments définis ci-dessus et la susdite substance fluorescente, la substance fluorescente étant telle que soit elle est excitable à la longueur d'onde d'émission de la GFP ou de l'un des susdits variants, ou de l'un des susdits fragments, soit elle émet à la longueur d'excitation de la GFP, ou de l'un des susdits variants, ou de l'un des susdits fragments, ou

entre la GFP ou l'un de ses variants définis ci-dessus, ou l'un des fragments définis ci-dessus, et la susdite molécule susceptible d'absorber la lumière émise par la protéine fluorescente.

35 L'invention a également pour objet l'utilisation d'une protéine G marquée par un marqueur constitué:

- soit par une molécule susceptible d'absorber la lumière émise par la protéine fluorescente,

5 - soit par une substance fluorescente pouvant être notamment choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est supérieur à environ $14000M^{-1}cm^{-1}$ et le rendement quantique de fluorescence est supérieur à environ 0,38, notamment choisie parmi :

- la protéine fluorescente verte (GFP), ou

10 - des variants dérivés de la GFP, par addition, délétion ou substitution d'un ou plusieurs acides aminés, sous réserve que ces variants conservent la propriété de fluorescence,

- ou des fragments de la GFP, ou de fragments des susdits variants, sous réserve que ces fragments conservent la propriété de fluorescence,

15 pour la détection et la quantification d'interactions non covalentes entre une protéine cible constituée par un récepteur couplé aux protéines G et la susdite protéine G, en vue d'identifier les molécules biologiquement actives vis-à-vis du récepteur, susceptibles de former une interaction non covalente réversible avec ledit récepteur, ledit récepteur étant marqué par voie génétique par une protéine fluorescente choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires dont le coefficient d'extinction moléculaire est supérieur à environ $14000M^{-1}cm^{-1}$ et le rendement quantique de fluorescence est supérieur à environ 0,38, notamment choisie parmi :

20 - la protéine fluorescente verte (GFP), ou

25 - des variants dérivés de la GFP, par addition, délétion ou substitution d'un ou plusieurs acides aminés, sous réserve que ces variants conservent la propriété de fluorescence,

- ou des fragments de la GFP, ou de fragments des susdits variants, sous réserve que ces fragments conservent la propriété de fluorescence,

30 cette détection et quantification ayant lieu par transfert d'énergie de fluorescence:

35 . entre la GFP ou l'un des variants définis ci-dessus, ou l'un des fragments définis ci-dessus et la susdite substance fluorescente, la substance fluorescente étant telle que soit elle est excitable à la longueur d'onde d'émission de la GFP ou de l'un des susdits variants, ou de l'un des susdits fragments, soit elle émet à la longueur d'excitation de la GFP, ou de l'un des susdits variants, ou de l'un des susdits fragments, ou

entre la GFP ou l'un de ses variants définis ci-dessus, ou l'un des fragments définis ci-dessus, et la susdite molécule susceptible d'absorber la lumière émise par la protéine fluorescente.

Selon un mode de réalisation avantageux du procédé de l'invention, la protéine fluorescente est choisie parmi EGFP, ECFP, EYFP, ou leurs mutants pour lesquels le coefficient d'extinction moléculaire est supérieur à environ $14000\text{ M}^{-1}\text{cm}^{-1}$ et le rendement quantique est supérieur à environ 0,38.

Selon un mode de réalisation avantageux le récepteur est choisi parmi :

- les récepteurs couplés aux protéines G, notamment ceux décrits dans Supplément Trends in Pharmacological Sciences, 1997 (*Receptor and ion Channel Nomenclature*),

- les séquences codant pour des récepteurs couplés aux protéines G putatifs dont les molécules biologiquement actives vis-à-vis d'eux-mêmes sont à identifier, notamment choisies parmi les séquences de récepteur orphelins disponibles dans les banques de séquence Genbank, EMBL et banques affiliées.

Selon un autre mode de réalisation avantageux, la protéine G est choisie parmi les protéines G décrites dans Journal of Receptor Research, vol 13, pp19-26, 1993 ou Angewandte Chemie, ed. Engl. vol 34, pp1406-1419, 1995.

L'invention concerne également un procédé de détection et de quantification d'interactions non covalentes entre une protéine cible constituée par un récepteur couplé aux protéines G et une protéine G, en vue d'identifier les molécules biologiquement actives vis-à-vis du récepteur, susceptibles de former une interaction non covalente réversible avec ledit récepteur, caractérisé en ce que :

- on prépare des cellules soit des fragments de cellules exprimant une séquence d'ADN comprenant le gène codant pour une protéine fluorescente fusionnée avec le gène du récepteur couplé aux protéines G, la fusion entre le gène de la protéine fluorescente et le gène du susdit récepteur étant telle que les propriétés du récepteur ne sont pas modifiées par la présence de la protéine fluorescente, à savoir :

- * l'interaction entre le récepteur et la protéine G n'est pas modifiée,
 - * l'interaction entre le récepteur et la molécule biologiquement active n'est pas modifiée,
 - * la fonction de transduction de la réponse n'est pas modifiée,
- la protéine fluorescente étant choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires dont le coefficient d'extinction moléculaire est supérieur à environ 14000M^{-1}

cm^{-1} et le rendement quantique de fluorescence est supérieur à environ 0,38, notamment choisie parmi :

- la protéine fluorescente verte (GFP), ou
- des variants dérivés de la GFP, par addition, délétion ou substitution d'un ou plusieurs acides aminés, sous réserve que ces variants conservent la propriété de fluorescence,
- ou des fragments de la GFP, ou de fragments des susdits variants, sous réserve que ces fragments conservent la propriété de fluorescence,

la protéine G étant marquée par un marqueur constitué :

- soit par une molécule susceptible d'absorber la lumière émise par la protéine fluorescente,
- soit par une substance fluorescente,
- la protéine fluorescente et le susdit marqueur étant tels qu'ils peuvent transférer l'énergie de l'un à l'autre, la protéine fluorescente pouvant être donneur d'énergie ou le susdit marqueur pouvant être donneur d'énergie,

l'interaction entre le récepteur marqué par la protéine fluorescente et la protéine G marquée par un marqueur susdéfini étant détectée par transfert d'énergie de fluorescence.

L'invention concerne également un procédé d'identification et éventuellement de quantification d'interactions entre un récepteur et une molécule non fluorescente biologiquement active vis-à-vis dudit récepteur, susceptibles de former une interaction non covalente réversible avec ledit récepteur, par mise en œuvre du procédé défini ci-dessus, dans lequel une molécule non fluorescente biologiquement active est ajoutée à des cellules, ou fragments de cellules, exprimant l'ADN codant pour le récepteur marqué par la protéine fluorescente et pour la protéine G marquée par le marqueur, caractérisé en ce que :

- une molécule non fluorescente biologiquement active et agoniste déclenche une transduction de signal détectée par variation de transfert d'énergie entre le récepteur marqué par la protéine fluorescente et la protéine G marquée par le marqueur ;
- une molécule non fluorescente biologiquement active antagoniste inhibe la transduction de signal provoquée par un agoniste et détectée par variation de transfert d'énergie de fluorescence entre le récepteur marqué par la protéine fluorescente et la protéine G marquée par le marqueur.

L'invention a pour objet des cellules ou fragment de cellules contenant une séquence d'ADN comprenant le gène codant pour une protéine fluorescente fusionnée avec le gène d'une protéine cible, la protéine fluorescente étant

choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est supérieur à environ $14000 M^{-1} cm^{-1}$ et le rendement quantique de fluorescence est supérieur à environ 0,38, la fusion entre le gène de la protéine fluorescente et le gène de la susdite protéine cible étant telle que

- * les propriétés de la protéine cible ne sont pas modifiées par la présence de la protéine fluorescente, à savoir
- * l'interaction entre la protéine cible et le ligand n'est pas modifiée,
- * la fonction de transduction de la réponse n'est pas modifiée,

sous réserve que :

- * lorsque la protéine cible est le récepteur des glucocorticoïdes de rat fusionné en N-terminal avec successivement une séquence de purification comportant 6 histidines, un épitope hémagglutinine et une protéine fluorescente et est exprimée dans la lignée cellulaire 1471.1 (Htun *et al.* 1996, PNAS 93:4845-4850), la protéine fluorescente est différente de la GFP (768 paires de bases du plasmide TU65 avec la mutation S65T)(Chalfie *et al.* 1994, Science 263:802-805, avec la mutation S65T),
- * lorsque la protéine cible est le récepteur humain des glucocorticoïdes tronqué de ses 131 premiers amino acides, fusionné en C-terminal d'une protéine fluorescente dans les sites Sal I et BamH I et est exprimée dans les cellules Cos-1 (Ogawa *et al.* 1995, PNAS 92:11899-11903), ladite protéine fluorescente est différente de la GFP telle que décrite dans l'article de Inouye S. et Tsuji, F. I., 1994, Febs Letters, 341:277-280,
- * lorsque la protéine cible est la sous-unité NMDA R1 de rat exprimé dans les cellules HEK 293 (selon Marshall *et al.* 1995, Neuron 14:211-215) fusionnée en C-terminal avec une protéine fluorescente, la protéine fluorescente est différente de celle constituée par les acides aminés 2-238 de la GFP sauvage (Chalfie *et al.* 1994, Science 263:802-805),
- * lorsque la protéine cible est un récepteur ou un fragment de récepteur de messagers secondaires intracellulaires, la protéine fluorescente est différente de la GFP et de ses dérivés (WO96/23898).

L'invention a également pour objet une trousse ou nécessaire pour la détection et la quantification d'interactions non covalentes entre une protéine

cible marquée par une protéine fluorescente et l'un de ses ligands marqué par un marqueur constitué :

- soit par une molécule susceptible d'absorber la lumière émise par la protéine fluorescente,

5 - soit par une substance fluorescente,

laquelle protéine fluorescente est choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est supérieur à environ $14000M^{-1}cm^{-1}$ et le rendement quantique de fluorescence est supérieur à environ 0,38, notamment choisie 10 parmi :

- la protéine fluorescente verte (GFP), ou

- des variants dérivés de la GFP, par addition, délétion ou substitution d'un ou plusieurs acides aminés, sous réserve que ces variants conservent la propriété de fluorescence,

15 - ou des fragments de la GFP, ou de fragments des susdits variants, sous réserve que ces fragments conservent la propriété de fluorescence et son ligand marqué par une substance fluorescente, ladite trousse comprenant :

- la protéine cible fusionnée avec une protéine fluorescente ou une lignée cellulaire stable susceptible d'exprimer la protéine fusionnée avec une protéine fluorescente ou un plasmide contenant la séquence nucléique codant pour ladite protéine cible fusionnée avec une protéine fluorescente telle que définie ci-dessus,

- le ligand marqué par le susdit marqueur,

20 - les tampons et milieux nécessaires au transfert d'énergie entre la susdite protéine et le susdit ligand.

L'invention a pour objet une trousse ou nécessaire pour la détection et la quantification d'interactions non covalentes entre une protéine cible marquée par une protéine fluorescente (n° 1) et l'un de ses ligand marqué par une substance fluorescente correspondant à une protéine fluorescente (n° 2), la protéine fluorescente (n° 1) étant choisie parmi la protéine fluorescente EYFP ou EGFP et le ligand étant marqué par la protéine fluorescente (n° 2) ECFP ou la protéine fluorescente

30 (n° 1) étant ECFP et le ligand étant marqué par la protéine fluorescente (n° 2) EYFP ou EGFP, ladite trousse comprenant :

- soit un plasmide contenant un acide nucléique codant pour la protéine cible fusionnée avec une protéine fluorescente (n° 1), et

* un plasmide contenant un acide nucléique codant pour le ligand fusionné avec une protéine fluorescente (n° 2), ou

- * un ligand fusionné avec une protéine fluorescente (n° 2), obtenu par voie recombinante et purifié,
 - soit une lignée cellulaire stable susceptible d'exprimer la protéine cible fusionnée avec une protéine fluorescente (n° 1), et
- 5
 - * une lignée cellulaire stable susceptible d'exprimer le ligand fusionné avec une protéine fluorescente (n° 2), ou
 - * un ligand fusionné avec une protéine fluorescente (n° 2), obtenu par voie recombinante et purifié,
- les tampons et milieux nécessaires au transfert d'énergie entre la susdite protéine et le susdit ligand.

10 L'invention a pour objet une trousse ou nécessaire pour la détection et la quantification d'interactions non covalentes entre une protéine cible constituée par un récepteur couplé à la protéine G marquée par une protéine fluorescente (n° 1) et la protéine G marquée par une substance fluorescente correspondant à une protéine fluorescente (n° 2), la protéine fluorescente (n° 1) étant choisie parmi la protéine fluorescente EYFP ou EGFP et la protéine G étant marquée par la protéine fluorescente (n° 2) ECFP ou la protéine fluorescente (n° 1) étant ECFP et la protéine G étant marquée par la protéine fluorescente (n° 2) EYFP ou EGFP, ladite trousse comprenant :

- 15
 - soit un plasmide contenant un acide nucléique codant pour le récepteur fusionné avec une protéine fluorescente (n° 1), et
- 20
 - * un plasmide contenant un acide nucléique codant pour la protéine G fusionnée avec une protéine fluorescente (n° 2), ou
 - * la protéine G fusionnée avec une protéine fluorescente (n° 2), obtenue par voie recombinante et purifiée,
- 25 - soit une lignée cellulaire stable susceptible d'exprimer le récepteur fusionné avec une protéine fluorescente (n° 1), et
- 30
 - * une lignée cellulaire stable susceptible d'exprimer la protéine G fusionnée avec une protéine fluorescente (n° 2), ou
 - * la protéine G fusionnée avec une protéine fluorescente (n° 2), obtenue par voie recombinante et purifiée,
- les tampons et milieux nécessaires au transfert d'énergie entre le susdit récepteur et la susdite protéine G.

DESCRIPTION DETAILLEE DE L'INVENTION

5 Dans sa réalisation préférée, le développement de l'invention utilise l'ADNc codant pour la protéine fluorescente verte (Prasher *et al.* Gene 1992, 111:229-233; GenBank Accession N° M62653) de la méduse *Aequorea victoria*, préférentiellement les mutants GFPUV, RSGFP et BFP de cette protéine fluorescente optimisés pour leur expression dans les organismes hôtes préférés, 10 les cellules mammifères.

L'ADNc peut être modifié pour coder pour un variant dans lequel un ou plusieurs acides aminés sont substitués, insérés ou déletés afin de permettre sa fusion, en N ou en C-terminal avec le gène codant pour une protéine cible.

La protéine cible, préférentiellement un récepteur, peut être choisie parmi:

15 1) les récepteurs de neurotransmetteurs couplés à des protéines G structurellement reliés aux récepteurs adrénériques et récepteurs métabotropiques du glutamate tels que présentés dans la liste actualisées annuellement et éditée comme supplément sous le nom: "Receptor and Ion Channel Nomenclature" par Elsevier Trends journals, dans Trends in Pharmacological Sciences.

20 2) les récepteurs-canaux structurellement reliés aux récepteurs nicotiniques, au récepteur de glutamate et au récepteur de l'ATP, tels que présentés dans la liste actualisées annuellement et éditée comme supplément sous le nom: "Receptor and Ion Channel Nomenclature" par Elsevier Trends journals dans Trends in Pharmacological Sciences.

25 3) les récepteurs nucléaires possédant un domaine d'interaction avec l'ADN structurellement reliés au récepteur des stéroïdes (Mangelsdorf *et al.* 1995, Cell, 83:835-839, Wurtz, J.L. *et al.* 1996, Nature Struct. Biol. 3:206).

30 4) les récepteurs de la membrane plasmique à activité tyrosine kinase structurellement reliés au récepteur de l'insuline (Yarden, Y. and Ullrich, A. 1988, Biochemistry 27:3113-3119).

35 5) les récepteurs membranaires couplés aux protéines tyrosine kinases (STATs, TYK2, Jak) structurellement reliés au récepteur de l'interféron γ (Brisco, J. *et al.* 1996, Phylos. Trans. R. Soc. Lond. B. Biol. Sci. 351:167-171; Ihle, J.N. 1995, Nature 377:591-594).

Dans le cas où la fusion est effectuée entre la EGFP et un récepteur couplé aux protéines G (groupe 1), la fusion peut être notamment effectuée:

- 1) du côté N-terminal du récepteur, et donc du côté C-terminal de la EGFP,
- 2) du côté C-terminal du récepteur et donc du côté N-terminal de la EGFP,
- 5 3) dans la séquence du récepteur, en particulier dans la première ou la troisième boucle intracellulaire, éventuellement en introduisant une ou plusieurs copies d'une séquence espaceur, notamment -GGGGS-.

Dans le cas où la fusion est effectuée entre la EGFP et un récepteur-canal (groupe 2), la fusion peut être notamment effectuée:

- 10 1) dans la région homologue à la "région immunogénique majeure" de la sous-unité α du récepteur nicotinique de Torpille (résidus 67-76), éventuellement en introduisant une ou plusieurs copies d'une séquence espaceur, notamment -GGGGS-.

Dans le cas où la fusion est effectuée entre la EGFP et un récepteur nucléaire (groupe 3), la fusion peut être notamment effectuée:

- 15 1) du côté N-terminal du récepteur, et donc du côté C-terminal de la EGFP,
- 2) du côté N-terminal du récepteur, tronqué dans sa partie N-terminale en amont du domaine de liaison à l'ADN, et donc du côté C-terminal de la EGFP.

20 Dans le cas où la fusion est effectuée entre la EGFP et un récepteur soit à activité tyrosine kinase, soit couplé à une tyrosine kinase (groupes 4 et 5), la fusion peut être notamment effectuée:

- 25 1) du côté N-terminal du récepteur, et donc du côté C-terminal de la EGFP,

Tout gène codant pour une protéine fluorescente, notamment la GFP, couplée à un récepteur, et dérivant d'organismes exprimant la GFP ou des protéines similaires pourraient être utilisés dans cette invention.

30 Les séquences d'ADN codant pour la GFP et les protéines cibles, notamment des récepteurs, peuvent être d'origine génomique ou être des ADNc, et peuvent être obtenus à partir de l'ADN de n'importe quelle espèce animale ou végétale, eucaryote ou procaryote, par exemple en préparant des banques génomiques ou des banques d'ADNc et en criblant ces banques pour identifier les séquences codantes par hybridation avec des sondes oligonucléotidiques par les techniques standard (Current Protocols in Molecular Biology, op. cit.).

35 Les constructions d'ADN codant pour la GFP et les protéines cibles peuvent aussi être obtenues par synthèse totale par les méthodes standards, notamment la méthode des phosphoramidites (Beaucage and caruthers, Tett.

Lett. 1981, 22:1859-1869) et l'emploi d'appareils de synthèse d'ADN automatisés, le polynucléotide obtenus étant ensuite purifiés, ligués et clonés dans les vecteurs appropriés. Pour la plupart des applications, les gènes codant pour la GFP et les protéines cibles seront préférentiellement obtenus par ciblage de banques, tandis que les bras espaces ainsi que les oligonucléotides requis pour la mutagenèse seront préférentiellement obtenus par synthèse.

Les constructions d'ADN pourront être de nature mixte, synthétique et génomique, par ligation de fragments synthétiques avec des élément d'ADN génomique, selon des procédures standard (Current Protocols in Molecular Biology, op.cit.).

Les constructions d'ADN peuvent aussi être obtenues par PCR (polymerase chain reaction) en employant des amores spécifiques, comme par exemple décrit dans PCR protocol 1990, Academic press, San Diego, California, USA.

Enfin, les constructions d'ADN peuvent être modifiées par d'autres méthodes incluant par exemple des réactions chimiques, de la mutagenèse aléatoire ou dirigée, par insertion, délétion ou substitution de nucléotides, ces modifications pouvant altérer des propriétés de l'une ou l'autre protéine, notamment, la GFP et les protéines cibles.

Les constructions d'ADN peuvent être insérées dans un vecteur recombinant. Ce vecteur peut être n'importe quel vecteur approprié pour les procédures employées avec des vecteurs recombinants. Le choix du vecteur sera souvent effectué en fonction de la cellules hôte dans laquelle a construction d'ADN voudra être introduite. Le vecteur pourra ainsi être un vecteur capable de se répliquer de manière autonome, c'est-à-dire extrachromosomal, et indépendante de la réplication chromosomale, par exemple un plasmide. Alternativement, le vecteur pourra être élaboré de manière à intégrer tout ou partie de l'ADN qu'il contient dans le génome de la cellules hôte, et se répliquera en même temps que le(s) chromosome(s) dans lequel il sera intégré.

Le vecteur est préférentiellement un vecteur d'expression dans lequel la GFP fusionnée à la protéine cible ou la GFP fusionnée au ligand est sous le contrôle d'autres segments d'ADN requis pour la transcription. En général, le vecteur d'expression dérive d'ADN plasmidique ou viral ou peut contenir des éléments de l'un et de l'autre.

Le terme "sous le contrôle" indique que les segments d'ADN sont disposés sur le vecteur de manière à fonctionner de concert pour servir l'objectif voulu, par exemple, la transcription est initiée dans le promoteur et se poursuit

tout au long de la séquence codant pour la protéine cible fusionnée à la GFP ou le ligand fusionné à la GFP.

Le promoteur peut être n'importe quelle séquence d'ADN susceptible de promouvoir une activité transcriptionnelle dans la cellule hôte choisie et peut être dérivé de gène homologues ou hétérologues à la cellule hôte.

Des exemples du promoteurs convenant pour l'expression de la protéine cible fusionnée avec le GFP ou du ligand fusionné avec la GFP dans des cellules mammifères sont le promoteur du virus simien SV40 (Subramani *et al.* 1981, Mol Cell. Biol. 1:854-864), le promoteur du virus du sarcome de Rous (RSV), le promoteur du cytomegalovirus (CMV) ou le promoteur tardif majeur de l'adénovirus (AdMLP).

Exemples de promoteurs pour cellules d'insectes :

Promoteur de la polyhédrine (US 4,745,051; Vasuvedan *et al.* 1992, FEBS Lett. 311:7-11) le promoteur P10 (Vlack *et al.* 1988, J. Gen. Virol. 69:765-776) le promoteur du gène précoce 1 du baculovirus (US 5,155,037; US 5,162,222).

Exemples de promoteurs pour levures :

Promoteurs des gènes de la glycolyse (Hitzeman *et al.* J. Biol. Chem. 1980, 255:12073-12080; Alber et Kawasaki, J. Mol. Appl. Gen. 1982, 1:419-434) des gènes des alcool deshydrogénases (Young *et al.* dans Genetic Engineering of microorganisms for chemicals (Hollaender *et al.* eds), Plenum Press, NY 1982).

Exemples de promoteurs pour bactéries:

Des exemples de promoteurs pour l'expression dans la bactérie peuvent être des promoteurs constitutifs comme le promoteur de la polymérase T7, ou des promoteurs inducibles comme par exemple le promoteur pL du phage lambda (Current Protocols in Molecular Biology, op.cit.).

Exemples de promoteurs pour champignons filamenteux

Les promoteurs utilisables sont par exemple le promoteur ADH3 (McKnight *et al.* EMBO J. 1985, 4:2093-2099) ou le promoteur tpiA. D'autres promoteurs utiles peuvent être dérivés des gènes codant pour l'aspartate protéinase de *Rhizomucor miehei*, l'alpha-amylase neutre de *Aspergillus niger*, de l'acétamidase de *Aspergillus nidulans*, de la TAKA amylase de *Aspergillus oryzae* ou le promoteur de la glucoamylase de *Aspergillus awamori*.

Le vecteur pourra par ailleurs contenir :

- des séquences de polyadénylation, comme par exemple celles de SV40 ou de la région Elb 5 de l'adénovirus,

5 - des séquences activatrices (enhancer) de la transcription (l'activateur de SV40),

- des séquences de réplication comme par exemple les séquences de réplication de SV40 ou du virus Epstein Barr, pour les cellules mammifères ou l'origine et les gènes de réplication REP 1-3 du plasmide 2 μ , pour les levures,

10 - des marqueurs de sélection, à savoir des gènes conférant une résistance à un antibiotique (néomycine, zéocine, hygromycine, ampicilline, kanamycine, tétracycline, chloramphénicol...) ou permettant la compensation d'un défaut (gène codant pour la dihydrofolate réductase permettant la résistance au méthotrexate, ou gène TPI de *S. pombe* décrit par Russell, 1985, Gene, 40:125-130).

15 La cellule hôte peut être n'importe quelle cellule capable d'exprimer la construction d'ADN insérée dans un vecteur approprié.

Les cellules peuvent être notamment des bactéries, des levures, des champignons et des cellules eucaryotes supérieures comme par exemple des cellules mammifères.

20 Des exemples de cellules bactériennes capables d'exprimer les constructions d'ADN sont :

- des bactéries grampositives telles que les souches de *Bacillus* comme *B. subtilis*, *B. licheniformis*, *B. lentus*, *B. brevis*, *B. stearothermophilus*, *B. thuringiensis* ou des souches de *Streptomyces* comme *S. lividans*, *S. murinus*,

25 - des bactéries gramnégatives telles *Escherichia coli*.

La transformation des bactéries peut être effectuée par transformation protoplastique ou par transformation de bactéries compétentes (Current Protocols in Molecular Biology, op.cit.).

30 Exemples de cellules eucaryotes:

Lignées cellulaires HEK 293, HeLa, cultures primaires, cellules COS (e.g. ATCC CRL 1650), BHK (e.g. ATCC CRL 1632) CHO (e.g. ATCC CCL 61).

35 Les méthodes d'introduction de l'ADN dans ces cellules (transfection, lipofection, électroporation etc. sont décrites dans Current Protocols in Molecular Biology, op.cit.

Exemples de cellules de levures :

Saccharomyces, S. cerevisiae, S. kluyveri,

Kluiveromyces, K. lactis,

Hansenula, H. polymorpha,

5 *Pichia, P. pastoris,*

transformées par introduction d'ADN hétérologue selon les protocoles décrits dans Current Protocols in Molecular Biology, op.cit.

Les cellules transformées sont sélectionnées par un phénotype déterminé par un marqueur de résistance, généralement à une drogue, ou par leur capacité à proliférer en l'absence d'un nutriment particulier.

10

Exemples de champignons filamenteux:

Les souches *Aspergillus* (*A. oryzae*, *A. nidulans*, *A. niger*), *Neurospora*, *Fusarium*, *Trichoderma*. L'utilisation de *Aspergillus* pour l'expression de protéines est décrite dans EP 272 277 ou EP 230 023 ou EP 184 438.

15

Exemples de cellules d'insectes :

On peut citer les lignées de *Lepidoptera* e.g. *Spodoptera frugiperda* (Sf9) ou *Trichoplusia ni* (Tni). Les méthodes de transformation (infection en particulier) sont décrites dans Current Protocols in Molecular Biology (op.cit.).

20

LES LIGANDS

Les ligands interagissant avec la protéine cible peuvent être de n'importe quelle origine (naturelle, synthétique, semi-synthétique, recombinante), et de n'importe quelle structure (chimique, peptidique, protéique). Ils peuvent être naturellement fluorescents (ou porteurs d'un chromophore) ou peuvent nécessiter soit une réaction chimique permettant le greffage d'un groupe fluorescent (ou d'un précurseur de groupe fluorescent) ou d'un chromophore, soit une construction d'ADN conduisant à la fusion du ligand avec la GFP et permettant l'expression du ligand ainsi rendu fluorescent.

25

30

Des exemples de réactions chimiques sont:

- couplage d'amines ou de thiols avec des réactifs de type halogénure d'alkyle, halogénures d'aryles, halogénures d'acyles, halogénures d'acides, le groupe isothiocyanate, le groupe maléimide, les époxydes, dans un solvant organique en présence d'une base ou en milieu aqueux,

35

- couplage d'acides avec des amines activées par des groupes tels que les succinimides.

Selon le procédé de l'invention, la fluorescence des cellules transformées peut être mesurée dans un spectrofluorimètre à l'aide duquel les propriétés spectrales des cellules, en suspension ou adhérentes, peuvent être déterminées par l'acquisition de leurs spectres d'excitation et d'émission. Les interactions avec le ligand fluorescent sont ensuite détectées par les changements des spectres d'excitation et/ou d'émission du donneur et de l'accepteur d'énergie, et les ligands sont définis comme pharmacologiquement significatifs si leurs interactions avec la protéine cible sont inhibées par l'addition d'un excès de ligand non fluorescent empêchant l'interaction entre la protéine cible fluorescente et le ligand fluorescent.

DESCRIPTION DES FIGURES :

- La Figure 1 donne la séquence nucléotidique codante de la GFP sauvage (Prasher *et al.* 1992, Gene 111:229-233) de *Aequorea victoria*.

- La Figure 2 représente la fluorescence mesurée lors de l'expression de la construction pMT3-EGFP-C3-SP. En ordonnée, on a présenté la fluorescence exprimée en coups par seconde (cps) et en abscisse la longueur d'onde d'émission. La trace représentée est un spectre d'émission de différence entre la mesure du milieu de culture exprimant la EGFP et le milieu de culture n'exprimant pas la EGFP.

- La Figure 3 représente la construction d'ADN pCEP4-NK2R-RF1 (13,67 kb) et comporte les séquence codant pour le peptide signal du récepteur nicotinique alpha7 (alpha7 SP), la EGFP (EGFP C3-1) et le récepteur NK2R (NK2R ARNm) des tachykinines.

- La Figure 4a représente les spectres d'excitation (émission 540 nm) et d'émission (excitation 450 nm) de cellules HEK293 en suspension exprimant les constructions d'ADN pCEP4-NK2R et pCEP4-NK2R-RF1. En ordonnée, on a présenté la fluorescence exprimée en coups par seconde (cps) et en abscisse la longueur d'onde d'émission.

Les courbes en pointillés représentent les spectres d'excitation (émission 540 nm) de pCEP4-NK2R (traits fins) et pCEP4-NK2R-RF1 (traits gras).

Les courbes en trait plein représentent les spectres d'émission (excitation 450 nm) de pCEP4-NK2R (traits fins) et pCEP4-NK2R-RF1 (traits gras).

5 - La Figure 4b représente le spectre d'émission (excitation 450 nm) de cellules exprimant les récepteurs NK2R-WT (courbe n° 3), NK2R-RF1 (courbe n° 1) et NK2R-RF2 (courbe n° 2). En ordonnée, on a présenté la fluorescence exprimée en coups par seconde (cps) et en abscisse la longueur d'onde d'émission.

10 - La Figure 5a représente la liaison de l'antagoniste ^3H SR 48968 sur des cellules exprimant le récepteur NK2R sauvage. En ordonnée, on a représenté la quantité de ^3H SR 48968 liée (en désintégrations par minute;dpm) et en abscisse, la concentration de ^3H SR 48968 ajoutée dans l'échantillon. Les losanges gris et noirs représentent la liaison totale de ^3H SR 48968 dans deux expériences différentes, les croix et carrés gris représentent la liaison non spécifique déterminée en présence d'un excès de neurokinine A ($10 \mu\text{M}$). Les lignes correspondent aux courbes théoriques de liaison d'un ligand à son récepteur ou à des sites de liaison non spécifiques. Les valeurs d'affinité déterminées (KD) sont de 1.05 nM et 0.78 nM dans chacune des deux expériences. Les valeurs de liaison maximale (B_{\max}) sont toutes deux de $0.1 \text{ pMol}/25000$ cellules.

15 - La Figure 5b représente la liaison de l'antagoniste ^3H SR 48968 sur des cellules exprimant le récepteur fluorescent NK2R-RF1. En ordonnée, on a représenté la quantité de ^3H SR 48968 liée (en désintégrations par minute;dpm) et en abscisse, la concentration de ^3H SR 48968 ajoutée dans l'échantillon. Les losanges gris et noirs représentent la liaison totale de ^3H SR 48968 dans deux expériences différentes, les croix et carrés gris représentent la liaison non spécifique déterminée en présence d'un excès de neurokinine A ($10 \mu\text{M}$). Les lignes correspondent aux courbes théoriques de liaison d'un ligand à son récepteur ou à des sites de liaison non spécifiques. Les valeurs d'affinité déterminées (KD) sont de 0.8 nM et 0.92 nM dans chacune des deux expériences. Les valeurs de liaison maximale (B_{\max}) sont respectivement de 0.075 et $0.088 \text{ pMol}/25000$ cellules.

20 - La Figure 6a représente le test de fonctionnalité de réponse de libération de calcium intracellulaire (FURA 2) de cellules exprimant la construction d'ADN pCEP4-NK2R WT. En ordonnée, on a représenté la fluorescence à 510 nm (exprimée en coups par seconde), l'excitation ayant lieu à 340 nm et en abscisse, on a représenté le temps (seconde). Les réponses sont évoquées par

l'agoniste neurokinine A (NKA) et inhibées par l'antagoniste cyclo(-Gln-Trp-Phe-Gly-Leu-Met) (cyclopeptide). Dans l'expérience 1, on ajoute 10 nM de NKA. Dans l'expérience (2,3), on ajoute successivement 5 μ M de cyclopeptide (2), puis 10 nM de NKA (3).

5

- La Figure 6b représente le test de fonctionnalité de réponse de libération de calcium intracellulaire (FURA 2) de cellules exprimant la construction d'ADN pCEP4-NK2R-RF1. En ordonnée, on a représenté la fluorescence à 510 nm (exprimée en coups par seconde), l'excitation ayant lieu à 340 nm et en abscisse on a représenté le temps (seconde). Les réponses sont évoquées par l'agoniste neurokinine A (NKA) et inhibées par l'antagoniste cyclo(-Gln-Trp-Phe-Gly-Leu-Met) (cyclopeptide). Dans l'expérience 1, on ajoute 10 nM de NKA. Dans l'expérience (2,3), on ajoute successivement 5 μ M de cyclopeptide (2), puis 10 nM de NKA (3).

10
15

- La Figure 7 représente la purification du peptide NKA BO I par HPLC en phase inverse. En abscisse, on a représenté le temps (minute) et en ordonnée la densité optique (mV). La détection est effectuée à 2 longueurs d'onde : 219 nm (traits pointillés) et 530 nm (traits pleins).

20

Les pics identiques 1, 2 et 3 sont respectivement la NKA, le dérivé NKA-BODIPY, le réactif BODIPY-IA.

25

- La Figure 8a représente le déplacement de la liaison de 3 H SR48968 sur le récepteur NK2R-WT par NKA et NKA BO I. En ordonnée, on a représenté la radioactivité (exprimée en désintégrations par minute;dpm) et en abscisse la concentration en ligand (M).

Les cercles noirs représentent l'expérience de déplacement de 3 H SR48968 par la NKA-BO I ($KI = 16,15$ nM).

30

Les carrés blancs représentent l'expérience de déplacement de 3 H SR48968 par la NKA ($KI = 3,03$ nM).

35

- La Figure 8b représente le déplacement de la liaison de 3 H SR48968 sur le récepteur NK2R-RF1 par NKA (ronds noirs : ($KI = 2,1$ nM) et NKA-BO I (carrés blancs : ($KI = 16,08$ nM)). En ordonnée, on a représenté la radioactivité (exprimée en désintégrations par minute;dpm) et en abscisse la concentration en ligand.

- La Figure 9 représente l'expérience de transfert d'énergie entre le ligand fluorescent NKA-BO I et le récepteur fluorescent NK2R-RF1 déterminée à l'équilibre.

5 La courbe en trait plein représente l'émission de fluorescence (entre 490 et 600 nm) des cellules exprimant le récepteur fluorescent NK2R-RF1 (excitation 460 nm) (Etape 1).

La courbe en grands traits pointillés représente le même spectre d'émission de fluorescence après addition de 100 nM de NKA BO I (Etape 2).

10 La courbe en traits pointillés représente le spectre d'émission de fluorescence après addition, sur l'étape 2, de 10 μ M de NKA non fluorescente. Les mêmes résultats sont obtenus après addition de 10 μ M de cyclo(-Gln-Trp-Phe-Gly-Leu-Met) ou de SR 48968.

15 L'addition, en étape 2, de NKA non fluorescente au lieu de NKA BO I est sans effet sur le spectre d'émission de fluorescence. Celui-ci se superpose avec le spectre en trait plein.

En ordonnée, on a représenté la fluorescence en cps, en abscisse la longueur d'onde en nm.

20 - La Figure 10 représente la mesure en temps réel du transfert d'énergie entre 100 nM de NKA BO 1 et le récepteur NK2R-RF1 (1), et inhibition (2) par un excès de ligand non fluorescent (20 μ M de NKA ou de MEN 10,376 ou de cyclopeptide). En abscisse, on a représenté le temps en seconde et en ordonnée la fluorescence (coups par seconde) à 510 nm, l'excitation ayant lieu à 460 nm.

25 La mesure est effectuée avec une suspension de 10^6 cellules/ml.

30 - La Figure 11a représente la mesure en temps réel du transfert d'énergie entre 100 nM de NKA-TR et le récepteur NK2R-RF1 (1), et inhibition par un excès (2) de ligand non fluorescent (100 μ M de NKA ou de MEN 10,376 ou de cyclopeptide). En abscisse, on a représenté le temps en seconde et en ordonnée la fluorescence (coups par seconde) à 510 nm, l'excitation ayant lieu à 460 nm.

35 - La Figure 11b représente la mesure en temps réel du transfert d'énergie entre 100 nM de NKA-Eos et le récepteur NK2R-RF1 (1), et inhibition par un excès (2) de ligand non fluorescent (20 μ M de NKA ou de MEN 10,376 ou de cyclopeptide). En abscisse, on a représenté le temps en seconde et en ordonnée la fluorescence (coups par seconde) à 510 nm, l'excitation ayant lieu à 470 nm.

- 5 - La Figure 12 représente l'amplitude du signal de fluorescence en fonction de la concentration de NKA BO I ajoutée (0 à 256 nM) à une suspension de cellules (10^6 cellules/ml) exprimant le récepteur NK2R-RF1 et l'inhibition par un excès (A) de ligand non fluorescent (10 μ M de NK ou de MEN 10,376 ou de cyclopeptide). En abscisse, on a représenté le temps en seconde et en ordonnée la fluorescence (coups par seconde) à 510 nm, l'excitation ayant lieu à 450 nm.
- 10 - La Figure 13 représente le traitement des données de la Figure 12. En abscisse, on a représenté la concentration en NKA BO I (nM) et en ordonnée l'amplitude du signal de fluorescence exprimée en unités arbitraires (U.A.). La ligne en trait plein représente la courbe théorique de liaison de la NKA-BO I à ses sites et fournit une valeur d'affinité $KD = 24$ nM.
- 15 - La Figure 14 représente l'extinction de fluorescence mesurée à 510 nm (excitation à 460 nm) résultant de la liaison de 10 nM NKA-Bo sur le récepteur EGFP-NK2R exprimé à la surface de cellules HEK 293 (suspension à 10^6 cellules/ml) en absence, et en présence de concentrations croissantes (indiquées sur le graphe en nM) de neurokinine A (NKA). En ordonnée, on représente la fluorescence mesurée en coups par seconde (cps) et en abscisse, le temps en seconde.
- 20 - La Figure 15 correspond à l'expérience décrite dans la figure 14, effectuée en présence de neurokinine A (NKA) ou de SR48968. L'amplitude de fluorescence, normalisée par rapport à l'amplitude de fluorescence mesurée en absence de NKA ou SR48968, est reportée sur un graphe en fonction de la concentration de NKA ou SR48968 utilisée pour tracer une courbe de compétition pour la liaison de NKA-Bo contre la NKA ou le SR48968 sur le récepteur EGFP-NK2R. L'affinité du ligand non fluorescent est déterminée à la valeur d'inhibition de 50 % de la liaison de NKA-Bo qui correspond à l' IC_{50} du composé non fluorescent.
- 25 - La Figure 16 représente le spectre d'émission de cellules (à 10^6 cellules/ml) exprimant la protéine de fusion EGFP-Hind-CCR5. L'excitation est fixée à 460 nm. En ordonnées, on représente la fluorescence exprimée en coups par seconde (cps) et en abscisse, la longueur d'onde d'émission en nm.

- La Figure 17 représente un test de fonctionnalité de réponse de libération de calcium intracellulaire (INDO-1) de cellules exprimant la construction d'AND EGFP-Hind-CCR5. En ordonnée, on a représenté la fluorescence à 405 nm exprimée en coups par seconde, l'excitation ayant lieu à 335 nm, et en abscisse, on a représenté le temps en seconde. La réponse évoquée par 100 nM de la chimiokine RANTES est montrée.

- La Figure 18 représente la révélation de l'expression de la chimiokine fluorescente EGFP-RANTES par immunoblot à l'aide d'un anticorps anti-GFP. Les levures, ou leurs surnageants de culture, exprimant une protéine non fluorescente (le trypsinogène) ou la construction d'ADN codant pour EGFP-RANTES sont déposés sur gel dénaturant de polyacrylamide 12 %. Après électrophorèse, les protéines sont électrotransférées sur membrane de nitrocellulose et la présence de protéines comprenant la GFP est révélée par addition d'anticorps primaire anti-GFP de lapin puis d'anticorps secondaire anti-lapin couplé à la luciférase. Les bandes protéiques contenant la GFP sont révélées par chimioluminescence à l'aide du Kit ECL disponible chez Amersham.

La figure révèle la présence de EGFP-RANTES dans les pistes 3 et 6 (respectivement surnageant de culture de levures, et levures) exprimant la construction d'ADN-EGFP-RANTES (bande de 42 000 daltons). Les pistes 1 et 8 correspondent à la migration de marqueurs de poids moléculaires annotés 20, 30, 40 et 50 kilodaltons. La piste 2 correspond au surnageant de levures contrôle (trypsinogène), la piste 4 correspond au précipité de dialyse de EGFP-RANTES, la piste 5 correspond à la levure contrôle et la piste 7 correspond à la levure exprimant EGFP seule.

EXEMPLES

EXEMPLE 1 : CONSTRUCTIONS D'ADN COMPRENANT UNE FUSION ENTRE LA EGFP ET L'EXTREMITE AMINO TERMINALE DU RECEPTEUR NK2R DES TACHYKININES

I) Fusion de la EGFP avec un peptide signal :

L'ADNc codant pour le EGFP (Figure 1) est fusionné en phase avec la séquence codant pour le peptide signal de la sous-unité alpha 7 de poulet (Genbank accession N°: X522995) codant pour un récepteur nicotinique de l'acétylcholine comme suit:

On introduit sur les codons codant 1 à 9 de la EGFP un site de restriction pour l'endonucléase BsrG I à l'aide de l'oligonucléotide

5'GGTCGCCACCTGTACAAGAAGGGCGAGG3', des réactifs fournis dans le kit de mutagenèse RPN 1526 (Sculptor) fourni par la société Amersham, et de simple brin de pEGFP C3 préparé à partir du plasmide pEGFP C3 (Genbank Accession N° U57607) fourni par la société ClonTech. le mutant pEGFP C3-1 obtenu est séquencé, puis cloné en phase avec le peptide signal de alpha7 par ligation de deux fragments: le fragment BsrGI-Xho I de 5225 nt de pJL223 (Eiselé *et al.* 1993, Nature 366:479-483) et le fragment BsrG I - Xho I de 725 nt de pEGFP C3-1. Le plasmide pJL223 contient le gène de la protéine a7-V201-5HT3 entre les sites Not I et Xho I du vecteur pMT3 (Swick, A.G. *et al.* 1992, Proc. Natl. Acad. Sci. 89:1812-1816). la construction obtenue nommée pMT3-EGFP-C3-SP, est transitoirement exprimée dans des cellules HEK 293 (ATCC CRL 1573) après transfection au phosphate de calcium (Cheng et Okayama 1986), afin de vérifier que la construction est correcte. Le spectre d'émission de fluorescence (excitation 450 nm) des surnageants de culture de cellules exprimant pMT3-EGFP C3-SP ou de cellules non transfectées (concentrés 5 fois par centrifugation sur centrikon 10 (Amicon)) sont enregistrés. Sur la figure 2 qui montre la différence entre le spectre de cellules transfectées et non transfectées, on détecte nettement le pic d'émission de la EGFP, ce qui indique que la construction conduit effectivement à l'expression de EGFP sécrétée dans le milieu de culture.

II) Clonage du récepteur NK2R des tachykinines dans les vecteurs de mutagenèse KS et d'expression pCEP4 :

25 Le fragment Spe I - Hind III, 2997 nt, du plasmide prTKR1-1 (Pr. S. Nakanishi, Kyoto university, Japon, Biochem. Biophys. Res. Comm. 1989, 165:695-702), contenant l'ADNc codant pour le récepteur NK2R (Genbank accession N°: M31838) de rat, est ligué avec le fragment Spe I - Hind III, 3549 nt, du vecteur pBluescript KS (+) pour conduire au plasmide pKS NK2R.

30 Le fragment Not I - BsrG I, 1369 nt, du plasmide prTKR1-1 est ligué avec le fragment BsrG I - Xho I, 1663 nt, de pKS NK2R et le fragment Not I - Xho I, 10370 nt de pCEP4 pour conduire au plasmide pCEP4-NK2R.

III) Fusion de la EGFP entre le peptide signal de alpha7 et l'extrémité amino terminal du récepteur NK2R des tachykinines :

35 Le fragment Not I - Xho I, 816 nt, de pMT3-EGFP C3-SP est ligué avec le fragment Not I Xho I (digestion partielle Xho I), 12856 nt, de pCEP4-NK2R pour conduire à la construction pCEP4-NK2R-RF1 (figure 3).

EXEMPLE 2 : CONSTRUCTION D'ADN COMPRENANT LA FUSION DE LA GFP DANS LES BOUCLES INTRACELLULAIRES I1 ET I3 DU RECEPTEUR NK2R DES TACHYKININES

5 I) Introduction de sites de clonage de la EGFP dans la boucle i1 ou i3 du récepteur NK2R :

L'ADN simple brin de pKS NK2R est mutagénisé comme en 1 a) avec les oligonucléotides

10 i1: 5'CACGAGAGGATGTACAACCTCGAGCGCACAGTCACC3'
contenant les mutations pour les sites de clonage BsrG I et Xho I, permettant l'introduction de la EGFP entre les acides aminés 65 et 66 et

15 i3:
5'GTACCCAGACACCAGCTAGCAGATCTGAAGCTTCGCCATCAGGC3'
contenant les mutations pour les sites de clonage Nhe I, Bgl II et Hind III
permettant l'introduction de la EGFP entre les résidus 233 et 234 ou 233 et 238.

20 Les plasmides obtenus pKS NK2R-i1 et -i3 sont introduits dans des bactéries XL1Blue compétentes (transformation), et les échantillons d'ADN plasmidique isolés à partir des colonies résistantes à l'ampicilline sont criblés pour la présence des sites introduits respectivement par les mutations à l'aide des oligonucléotides i1 et i3.

II) Clonage de la EGFP dans les boucles i1 et i3 du récepteur NK2R des tachykinines :

25 - clonage dans la boucle i1: les fragments Hind III - BsrG I, et Hind III - Xho I, de pKS NK2R-i1 sont ligués avec le fragment BsrG I - Xho I, 725 nt, de pEGFP C3-1. L'insert de 3741 nt codant pour la protéine de fusion est ensuite excisé par les enzymes Spe I et Sal I et ligué avec le vecteur pCEP4 ouvert par les enzymes Nhe I et Xho I pour conduire à la construction d'ADN pCEP4 NK2R-RF2 ;

30 - clonage dans la boucle i3: deux constructions sont obtenues:

les fragments Nhe I - Bgl II, 744 nt, de pEGFP C3 sont ligués avec les fragments Not I - Nhe I et Bgl II - Not I de pKS-NK2R-i3, l'insert Spe I - Sal I de 3750 nt ainsi obtenu est ensuite cloné dans le vecteur pCEP4 entre les sites Nhe I et Xho I pour conduire à la construction d'ADN pCEP4-NK2R-RF3.

35 les fragments Nhe I - Hind III, 757 nt, de pEGFP C3 sont ligués avec les fragments Not I - Nhe I et Hind III - Not I de pKS-NK2R-i3, l'insert Spe I - Sal I de 3770 nt ainsi obtenu est ensuite cloné

dans le vecteur pCEP4 entre les sites Nhe I et Xho I pour conduire à la construction d'ADN pCEP4-NK2R-RF4.

5 EXEMPLE 3 : EXPRESSION DES PROTEINES RECOMBINANTES
ET CARACTERISATION FONCTIONNELLE

I) Expression :

Des cellules HEK 293 sont transfectées, par la méthode de la précipitation au phosphate de calcium (Chen & Okayama 1987, Mol. Cell. Biol. 7:2745-2752, Current Protocols in Molecular Biology, op.cit.), par les constructions 10 d'ADN pCEP4-NK2R-RF1, -RF2, -RF3 et -RF4. Des lignées stables sont établies par sélection des cellules transfectées résistantes à l'hygromycine (100 µg/ml, Clontech). Les cellules sont cultivées en présence d'hygromycine 15 100 µg/ml dans du milieu MEM (Gibco) supplémenté de 10% de sérum de veau foetal (Seromed) de Pénicilline (100 unités/ml), streptomycine (100 µg/ml) et glutamine (4 mM) (Methods in enzymology, Vol LVIII, 1979).

II) Mesures de l'expression et propriétés de liaison des récepteurs recombinants :

20 a) Par fluorimétrie :

Les expériences de fluorescence sont effectuées dans une cuve de 1 ml pourvue d'un système d'agitation magnétique et placée dans un spectrofluorimètre Fluorolog (SPEX) équipé d'une lampe Xe 450W (Osram) et de monochromateurs Spex 1680 0.22m (excitation) et Spex 1681 0.22 m 25 (émission). Les cellules ou les fragments de membrane sont mis en suspension dans du tampon physiologique: Hepes 10 mM, NaCl 137,5 mM, MgCl₂ 1,25 mM, CaCl₂ 1,25 mM, KCl 6 mM, glucose 5,6 mM, NaH₂PO₄ 0,4 mM, BSA 0,1% (w/v), pH 7,4.

Les cellules entières sont récoltées après traitement au Versène (PBS, 30 5 mM EDTA), centrifugées 5 min. à 1.000 g et sont resuspendues à une concentration de 250.000-1.000.000 cellules/ml dans le tampon physiologique.

La figure 4a montre les spectres d'excitation et émission enregistrés avec 35 des cellules transfectées par les constructions pCEP4-NK2R-WT et pCEP4-NK2R-RF1. Les spectres font clairement apparaître le signal de la EGFP et indiquent que ces constructions sont effectivement exprimées par les cellules HEK 293. La figure 4b montre les spectres émission de fluorescence de cellules exprimant le récepteur sauvage ainsi que les fusions -RF1 et RF2.

b) Par expérience de liaison :

L'apparition de sites de liaison des neurokinines à la surface des cellules est déterminée par liaison de 0.1 nM 125 I NKA ou 1 nM 3 H SR48968 dans du PBS sur des cellules entières en suspension (200-500.000 Cellules/ml, 250 μ l/essai).⁵ Après 30 min. d'incubation à température ambiante, les échantillons sont filtrés sur filtres GF/C (Whatmann) prétraités dans 1% de lait en poudre dans du PBS (KCl 2.7 mM, KH₂PO₄ 1.47 mM, NaCl 0.137 mM, Na₂HPO₄ 8,06 mM), rincés deux fois au PBS (4 ml), puis comptés dans un cocktail scintillant.

10 Les courbes de saturation de la liaison de 3 HSR 48 968 sont effectuées dans du tampon Hepes 10 mM, NaCl 137,5 mM, MgCl₂ 1,25 mM, CaCl₂ 1,25 mM, KCl 6 mM, glucose 5,6 mM, NaH₂PO₄ 0,4 mM, BSA 0,1% (w/v), pH 7,4 à 4°C pendant 1 h 30 en présence ou en absence de NKA 1 μ M final. Chaque échantillon de 500 μ l contient 25 000 cellules, et la fraction de ligand lié est inférieure à 10% de la quantité ajoutée dans chaque tube. La filtration est effectuée en fin d'incubation comme décrit plus haut. les figures 5a et 5b montrent les courbes de liaison (en dpm liés) obtenues avec des concentrations croissantes du ligand 3 HSR 48968 sur respectivement des cellules exprimant le récepteur sauvage (NK2R WT) ou la fusion N terminale avec la EGFP (NK2R-RF1).¹⁵ Le ligand exhibe la même affinité pour les deux récepteurs et le nombre de sites de liaisons exprimés à la surface des cellules est comparable.²⁰

III) Tests fonctionnels :

I) Mesure du calcium cytosolique à l'aide du FURA 2 :

25 Le chelateur fluorescent du calcium, FURA 2 acétylmethylester (Molecular probes), dissout dans du DMSO est dilué à la concentration de 3 μ M dans du tampon Hepes 10 mM, NaCl 137,5 mM, MgCl₂ 1,25 mM, CaCl₂ 1,25 mM, KCl 6 mM, glucose 5,6 mM, NaH₂PO₄ 0,4 mM, BSA 0,1% (w/v), pH 7,4. Le milieu de culture des cellules est aspiré puis remplacé par 7 à 10 ml de la solution de FURA 2 AM. Les cellules sont placées dans l'incubateur à CO₂ pendant 45 min. Le milieu est alors aspiré, remplacé par le tampon Hepes sans FURA et les cellules sont replacées dans l'incubateur à CO₂ pendant 30 15 min. supplémentaires. Le tampon est aspiré, puis les cellules sont récoltées après avoir été décrochées de la boîte par une solution de PBS, 5 mM EDTA pendant 5 min. Après centrifugation à 1000 x g pendant 5 min., le culot de cellules obtenu est resuspendu à raison de 1 x 10⁶ cellules/ml dans le tampon Hepes.³⁵

Les mesures de libération de calcium sont effectuées à 37°C, dans une cuve de spectrofluorimètre contenant par un barreau magnétique d'agitation. La longueur d'onde d'excitation est fixée à 340 nm et la longueur d'onde d'émission est fixée à 510 nm.

La figure 6, montre les résultats obtenus lors de la stimulation de cellules transfectées avec la construction codant pour le récepteur sauvage (6a) et de cellules transfectées par la construction pCEP4-NK2R-RF1 (6b). Les réponses évoquées par la neurokinine A sont inhibées par les antagonistes NK2R spécifiques SR48968 (Sanofi recherche), MEN 10,376 (Bachem) ou le cyclopeptide cyclo(-Gln-Trp-Phe-Gly-Leu-Met) (Bachem). La neurokinine A n'évoque aucune réponse sur des cellules HEK 293 non transfectées.

Préparation de fragments de membranes :

Les fragments de membrane sont préparés par homogénéisation des cellules à l'aide d'un homogénéiseur de tissu (potter ou ultraturax) en présence d'un mélange d'inhibiteurs de protéases, à 4°C. Les cellules sont centrifugées à 1000 x G pendant 5 min. puis resuspendues dans du tampon de préparation de membranes: Tris-HCl 50 mM, EDTA-Na 1 mM, DTT 10 mM, contenant le cocktail d'inhibiteurs de protéases, Complete (Boehringer). La suspension est homogénéisée à l'aide d'un potter à 4°C. Après centrifugation à 3000 x G pendant 10 min., le surnageant est collecté, le culot est repris dans le tampon ci-dessus, rehomogénéisé, et centrifugé à nouveau. Les deux surnageants obtenus sont rassemblés et centrifugés à 150 000 x g pendant 30 min. Les culots obtenus sont resuspendus dans du tampon de préparation de membranes à une concentration de 6-10 mg prot/ml.

EXAMPLE 4 : PREPARATION DE LIGANDS FLUORESCENTS

I) Préparation de neurokinine A fluorescente :

a) Groupe Bodipy 530/560 :

La NKA lyophilisée est dissoute dans de la DMF à 10 mM final.

A 50 µL de cette solution (0.5 µMol, 50 µL), on ajoute 10 µmol de NEt₃ (200 mM dans CH₃CN), soit 100 µL, puis 15 µL de BODIPY 530/550 IA [N-(4,4-difluoro-5,7-dimethyl-4-bora-3a,4a-diaza-s-indacene-3-propionyl)-N'-iodoacetylenediamine] dissout dans de la DMF à raison de 0.3 µMol par 10 µL. Le mélange est vortexé, puis laissé à température ambiante. Après 24 h, le produit de la réaction est purifié par HPLC (Gilson) sur une colonne de phase inverse Z5C8 25F (Zorbax) sur laquelle on développe un gradient linéaire de 10

à 95% solvant B en 60 minutes (A: H₂O 0.1% TFA; B: CH₃CN 0.1% TFA) avec un débit de 1 ml/min. Les longueurs d'ondes de détection sont fixées à 219 nm (peptide) et 530 nm (fluorophore). Le produit élué (NKA BO I) au temps 34 min. (Figure 7) est collecté, concentré par évaporation et resuspendu dans de la DMF.

5

b) Groupes coumarine et éosine :

Les conditions réactionnelles employées pour la marquage de la neurokinine A par les dérivés de coumarine (7-diéthylamino-3-((4'-iodoacétyl)amino)phényl)-4-méthylcoumarine) et d'éosine (éosine-5-iodoacétamide) sont les mêmes que en a). Les purifications sont effectuées sur colonne de phase inverse Z5C8 25F (Zorbax) sur laquelle on développe un gradient linéaire de 10 à 95% solvant B en 60 minutes (A: H₂O 0.1% TFA; B: CH₃CN 0.1% TFA) avec un débit de 1 ml/min. le dérivé coumarine de la NKA (NKA-Coum) est élué au temps 25 min., le dérivé éosine de la NKA (NKA-Eos) est élué au temps 27 min.

10

15

c) Groupe sulforhodamine 101 :

Le réactif utilisé est le chlorure d'acide de la sulforhodamine 101 (Aldrich, ou Texas Red commercialisé par Molecular Probes). Le protocole est le même que pour le greffage du BODIPY 530/550 IA. La purification est effectuée par HPLC sur une colonne de phase inverse Z5C8 25F (Zorbax) sur laquelle on développe un gradient linéaire de 10 à 95% solvant B en 60 minutes (A: H₂O 0.1% HFBA; B: CH₃CN 0.1% HFBA) avec un débit de 1 ml/min. Les longueurs d'ondes de détection sont fixées à 219 nm (peptide) et 590 nm (fluorophore). le produit élué (NKA TR) au temps 40-41 min. est collecté, concentré par évaporation et resuspendu dans de la DMF.

20

25

EXEMPLE 5 : DETECTION DE L'INTERACTION ENTRE LE RECEPTEUR NK2R FLUORESCENT ET SES LIGANDS FLUORESCENTS.

30

I) Par déplacement d'un radioligand (expérience de compétition) :

Les courbes de déplacement de la liaison de 3HSR 48968 par un ligand non radioactif sont effectuées dans du tampon Hepes 10 mM, NaCl 137,5 mM, MgCl₂ 1,25 mM, CaCl₂ 1,25 mM, KCl 6 mM, glucose 5,6 mM, NaH₂PO₄ 0,4 mM, BSA 0,1% (w/v), pH 7,4 à 4°C pendant 1 h 30 en présence ou en absence de NKA 1 µM final. Chaque échantillon de 500 µl contient 25 000 cellules, 1 nM de 3HSR 48968 et un ligand non radioactif à différentes

35

concentrations (de 1 nM à 10 μ M). L'incubation est effectuée pendant 1 heure à 4°C. La filtration est effectuée en fin d'incubation comme décrit en 3b).

La figure 8 montre les courbes de déplacement du radioligand 3HSR 48968 par la neurokinine A (NKA) et la NKA BO I sur respectivement des cellules HEK 293 exprimant le récepteur NK2R sauvage (construction pCEP4-NK2R, figure 8A) ou la fusion de la EGFP en N-terminal du récepteur NK2R (construction pCEP4-NK2R-RF1, figure 8b). Les points représentent la radioactivité liée (en dpm) en présence de concentrations variables de NKA ou de NKA BO I. Les valeurs d'affinité dérivées de ces courbes indiquent que la NKA (tout comme la NKA BO I) se lie avec la même affinité au récepteur sauvage et au récepteur rendu fluorescent. Les valeurs d'affinité sont pour la NKA: $K_I = 3$ nM sur le récepteur sauvage et 2,1 nM sur le récepteur fluorescent et pour la NKA BO I: $K_I = 16$ nM sur le récepteur sauvage et 16 nM sur le récepteur fluorescent. Les même mesures effectuées avec l'antagoniste MEN 10,376 conduisent aux valeurs d'affinité: $K_I = 53$ nM sur le sauvage et $K_I = 61$ nM sur le récepteur fluorescent.

II) Par transfert d'énergie :

a) interactions entre la NKA BO I et le récepteur NK2R-RF1 à l'équilibre:

Les cellules HEK293 exprimant le récepteur NK2R-RF1 sont placées en suspension à 1 000 000 cellules/ml dans le tampon physiologique Hepes (Hepes 10 mM, NaCl 137,5 mM, MgCl₂ 1,25 mM, CaCl₂ 1,25 mM, KCl 6 mM, glucose 5,6 mM, NaH₂PO₄ 0,4 mM, BSA 0,1% (w/v), pH 7,4). Les cellules sont ensuite placées dans une cuve de fluorescence à une concentration pouvant varier de 100 000 à 1 000 000 cellules/ml. Un spectre d'émission de ces cellules est enregistré entre 490 et 600 nm (excitation 460 nm) puis la NKA BO I est ajoutée à une concentration finale de 100 nM. On enregistre à nouveau le spectre d'émission de la solution. On ajoute ensuite un excès (1-10 μ M) de ligand non marqué spécifique du récepteur NK2R (NKA non marquée ou de SR 48968 (Sanofi recherche) ou du peptide cyclique, cyclo(-Gln-Trp-Phe-Gly-Leu-Met), et on enregistre à nouveau le spectre d'émission de la solution. La Figure 9 montre les trois spectres superposés sur lesquels on détecte nettement une extinction de la fluorescence de la EGFP (pic à 510 nm) lors de l'addition de NKA BO I et un retour à la fluorescence initiale lors de l'addition subséquente de ligand non marqué. on détecte aussi l'apparition d'un pic à 550-560 nm lors de l'addition du ligand fluorescent et une diminution de son intensité lors de l'addition de ligand non marqué. Ces résultats indiquent l'occurrence d'un transfert d'énergie entre le récepteur fluorescent et son ligand

fluorescent. Ce signal est renversable par addition d'un agent pharmacologique susceptible d'inhiber l'interaction entre le ligand et son récepteur. L'addition de NKA non fluorescente sur des cellules exprimant le récepteur fluorescent NK2R-RF1 est sans effet sur le spectre d'émission de fluorescence.

5

b) Interactions entre la NKA BO I et le récepteur NK2R-RF1 en temps réel :

10

Les mêmes enregistrements peuvent être effectués en temps réel. Pour cela, la longueur d'onde d'excitation est fixée à 460 nm et la longueur d'onde d'émission à 510 nm. L'addition de NKA BO I sur une suspension de cellules entraîne une diminution de l'intensité de fluorescence à 510 nm et l'addition subséquente d'un excès de ligand non marqué rétablit la fluorescence à sa valeur initiale (voir figure 10). La valeur finale de l'intensité de fluorescence peut être supérieure à la valeur initiale lorsque l'absorption du ligand fluorescent contribue au signal. Ceci est observé avec la NKA BO I pour des concentrations supérieures à 200 nM, lorsque la concentration en cellules est de l'ordre de 5×10^5 à 1×10^6 cellules/ml. Enfin, l'addition de NKA non fluorescente au temps 0 est sans effet sur l'intensité du signal mesuré à 510 nM.

15

c) Détection de l'interaction entre le récepteur NK2R-RF1 et les ligands NKA TR et NKA Eos :

20

La reproduction de l'expérience décrite en a) avec la NKA TR (figure 11a) ou la NKA Eos (figure 11b) indique que la détection d'un signal d'interaction récepteur NK2R-RF1 ligand fluorescent n'est pas restreinte au couple EGFP-BODIPY mais peut aussi être étendue aux couples EGFP-sulforhodamine 101 ou EGFP-Eosine.

25

d) Expérience de saturation de la liaison de NKA BO I à ses sites récepteurs :

30

Etant donné que l'amplitude du signal enregistré selon la description en 8b (figure 12) est proportionnel à la concentration de NKA BO I ajoutée, on peut effectuer une mesure de l'amplitude du signal spécifique en fonction de la concentration de ligand ajouté. La figure 12a montre la variation d'amplitude du signal lorsque la concentration de NKA BO I ajoutée varie de 0 à 256 nM, par addition successive d'aliquots d'une solution stock de 10^{-5} M sur une suspension de 1 ml de cellules à 10^6 cellules/ml dans le tampon physiologique Hepes. Le traitement des données expérimentales (figure 12b) permet d'extraire une valeur d'affinité de la NKA BO I ($K_D = 20-30$ nM) en excellent accord

35

avec les mesures effectuées par déplacement d'un radioligand ainsi qu'il est décrit dans l'exemple 5, I).

e) Expérience de compétition :

5 Les fragments de membrane contenant le récepteur NK2R-RF1 préparés selon le procédé décrit dans l'exemple 3, section préparation de fragments de membranes, sont dilués dans le tampon Hepes (Hepes 10 mM, NaCl 137,5 mM, MgCl₂ 1,25 mM, CaCl₂ 1,25 mM, KCl 6 mM, glucose 5,6 mM, NaH₂PO₄ 0,4 mM, BSA 0,1% (w/v), pH 7,4) à raison de 20 µL de membranes/ml. Des aliquots des membranes sont incubés en présence de NKA, MEN 10,376, de cyclo(-Gln-Trp-Phe-Gly-Leu-Met) ou de SR 48968 à différentes concentrations. Après 30 minutes d'incubation, les échantillons sont placés dans une cuve de spectrofluorimétrie et le signal engendré par addition de NKA BO I 25 nM est enregistré pendant 120 sec. à 510 nm, la longueur d'onde d'excitation ayant été fixée à 470 nm. L'amplitude du signal évoqué par l'interaction entre le récepteur NK2R-RF1 et son ligand NKA BO I étant proportionnel au nombre de sites de liaison accessibles au ligand fluorescent, on peut établir une courbe de déplacement de la NKA BO I par le ligand non fluorescent.

10 15

On peut aussi procéder par mélange simultané des fragments de membrane contenant le récepteur NK2R-RF1, le ligand NKA BO I et le ligand non marqué. Le mélange est ensuite incubé pendant le temps nécessaire à atteindre l'équilibre de liaison (30-60 min.) et la fluorescence est mesurée à 510 nm en fixant la longueur d'onde d'excitation à 470 nm. Les mêmes expériences peuvent être effectuées avec des cellules entières à la concentration de 250 000 à 25 1 000 000 de cellules/ml.

20 La figure 14 montre que l'amplitude du signal de fluorescence mesuré à 510 nm (excitation 460 nm) décroît lorsque la concentration de molécule non fluorescente présente dans l'essai augmente. La courbe d'inhibition de la liaison de NKA Bo I permet d'estimer l'affinité des molécules neurokinine A et SR48968 non fluorescentes.

25 30

La figure 15 correspond à l'expérience décrite dans la figure 14, effectuée en présence de neurokinine A (NKA) ou de SR48968.

a) Clonage :

La partie de pCEP4-NK2R-RF1 codant pour la protéine de fusion est amplifiée par PCR (Current Protocols in Molecular Biology, op.cit.) en utilisant les amores

5'GGAGAGTTCCAACTCGAGAAAAGAAAGAAGGGCGAGGAG3' et
5'GTCAGCTGTTCTGCGGCCGCTAACGCCTGGGCCTT3'

permettant 1) la production d'un fragment de 1868nt codant pour la totalité de la protéine de fusion NF2R-EGFP à l'exception du peptide signal et 2) le clonage en phase dans le vecteur d'expression de levure pPIC9 (invitrogen) avec la séquence codant pour le peptide signal du facteur alpha promoteur du gène AOX1. Les sites de clonage utilisés sont respectivement XHoI pour l'extrémité 5' du produit d'amplification et Not I pour son extrémité 3'

b) Expression :

Les levures sont transformées avec le plasmide pPIC9-NK2R-RF1 linéarisé (StuI ou SalI) et mises en culture sur milieu MD sans histidine, préparé selon les indications du manuel fourni avec le vecteur pPIC9 (inVitrogen). L'expression de la construction d'ADN introduite dans la cellules est induite par le Méthanol. Pour cela, les colonies sont propagées en milieu liquide (BMGY) pendant 24 h, puis transférées dans du milieu BMMY contenant 0.5% de méthanol permettant l'induction de l'expression de la construction d'ADN NK2R-RF1. Des aliquots de ces cultures sont prélevés et les clones exprimant la protéine fluorescente sont identifiés par mesure des spectres d'excitation et d'émission de la EGFP.

25

EXEMPLE 7 : CONSTRUCTION D'ADN CODANT POUR LE RECEPTEUR MUSCARINIQUE DE L'ACETYLCHOLINE FUSIONNE AVEC LA EGFP ET EXPRESSION DANS LES CELLULES MAMMIFERES.

30

I) Clonage : le fragment d'ADNc codant pour le récepteur humain muscarinique M1 (Genbank Accession N° X15263) est amplifié par PCR (Current Protocols in Molecular Biology, op.cit.) en utilisant les amores :

5'TTAGTTCTAAACTAGCGGCCGCACTAGTCCTCCATGAACACTTCAGC
35 CCCA3' et

5'CTTGAAACCTATAGCTAGCCTCGAGTCAGCATGGCGGGAGGG3'.

Le fragment de 1383 nt obtenu est clivé par les enzymes Not I (en 5') et Xho I (en 3') et ligué avec le vecteur KS ouvert par les mêmes enzymes

(fragment de 2888nt) pour conduire à la construction KS-hM1, ou ligué avec le vecteur pCEP4 ouvert par les mêmes enzymes pour conduire à la construction pCEP4-hM1.

La construction KS-hM1 est utilisée pour la production d'ADN simple
5 brin (Current Protocols in Molecular Biology, op.cit.) et l'obtention de mutants.

II) fusion de la EGFP en N-terminal du récepteur hM1.

L'oligonucléotide 5'CCTGCTGTCTCAGATCTCATCACCGTCC3' est
utilisé, conjointement avec les réactifs du kit de mutagenèse Sculptor
10 (Amersham) pour produire un mutant permettant la fusion en position 13 de la
séquence codante du récepteur hM1 grâce à l'introduction d'un site de
restriction pour l'enzyme Bgl II.

Le mutant obtenu est digéré par les enzymes Bgl II et Xho I et le fragment
15 de 1354 nt engendré, est ligué avec le fragment Not I - Bgl II, 812 nt, de
pCEP4-NK2R-RF1, et le vecteur pCEP4 ouvert par les enzymes Not I et Xho I,
pour conduire à la construction d'ADN pCEP4-hM1-RF1.

III) expression de la protéine de fusion hM1-RF1.

Le plasmide pCEP4-hM1-RF1 est introduit dans des cellules HEK 293 par
20 transfection au phosphate de calcium ou dans des cellules Cos 1 par
électroporation (Current Protocols in Molecular Biology, op.cit.).

L'expression de la protéine est détectée comme décrit plus haut.

IV) synthèse et purification du ligand muscarinique ABT-Bodipy.

25 L'ABT base (3-[2'-aminobenzhydryloxy]tropane) est mis en solution dans
la DMF (10 mM). 20 µl de cette solution (0.2 µMoles) sont mélangés avec 4 µl
d'une solution 100 mM de Bodipy -IA et laissés à température ambiante pendant
20 h. Le produit de la réaction est purifié par HPLC (Gilson) sur une colonne
30 de phase inverse Z5C8 25F (Zorbax) sur laquelle on développe un gradient
linéaire de 10 à 95% solvant B en 60 minutes (A: H₂O 0.1% TFA; B: CH₃CN
0.1% TFA) avec un débit de 1 ml/min. Les longueurs d'ondes de détection sont
fixées à 219 nm (peptide) et 530 nm (fluorophore). Le produit élué (ABT Bo) au
temps 34 min. (Figure 7) est collecté, concentré par évaporation et resuspendu
dans de la DMF.

35

V) fusion de la EYFP en N-terminal du récepteur hM1 et détection de
l'interaction avec la pirenzepine-bodipy 558/568.

Cette construction est réalisée de manière identique à la construction décrite au point II.

Les interactions de la protéine de fusion avec son ligand sont effectuées à l'aide du ligand pirenzépine-bodipy dans lequel le groupe fluorescent correspond au bodipy 558/568 (4,4-difluoro-5-(2-thiényl)-4-bora-3a,4a-diaza-s-indacene-3-propionic acid).

10 **EXEMPLE 8 : CONSTRUCTION D'ADN CODANT POUR DES RECEPTEURS NICOTINIQUES FLUORESCENTS**

a) Construction d'ADN codant pour la fusion de la EGFP dans la région MIR du récepteur a7-V201-5HT₃.

15 Le plasmide pJL223 (Eiselé *et al.* 1993, Nature 366:479-483) contient le gène de la protéine a7-V201-5HT₃ qui forme un récepteur canal activé par l'acétylcholine et la nicotine lors de son expression dans des ovocytes de Xénope. L'ADNc codant est compris entre les sites Not I et Xho I du vecteur pMT3 (Swick, A.G. *et al.* 1992, Proc. Natl. Acad. Sci. 89:1812-1816).

20 L'insert Not I - Xho I de 1424 nt est cloné entre les sites Not I et Xho I du vecteur Bluescript et le plasmide obtenu (KS 223) sert de matrice à la production d'ADN simple brin.

L'oligonucléotide

5'CAGATCATTAGTTGTACAGGAAAGATCTTGAGGATCCTGGAGTGAA
G3'

25 est utilisé pour introduire, sur KS223, les sites de restriction pour les enzymes BsrG I, Bgl II et BamH I dans la même phase que celle des sites identiques portés par le plasmide pEGFP-C3. La mutation est introduite au niveau d'une région du récepteur appelée MIR (Major Immunogenic Region, Barkas *et al.* 1987, Science, 235:77-80) entre les acides aminés 63 et 64.

30 Les fragments Not I-BsrG I (267 nt) et Bgl II-Xho I (1147 nt) de ce mutant de KS 223 sont liqués avec le fragment BsrG I-Bgl II (721 nt) de pEGFP-C3 et le vecteur pCEP4 ouvert par les enzymes Not I et Xho I, pour conduire à la construction d'ADN pCEP4-223-RF1.

35 Les fragments Not I-BsrG I (267 nt) et BamH I-Xho I (1138 nt) de ce mutant de KS 223 sont liqués avec le fragment BsrG I-BamH I (772 nt) de pEGFP-C3 et le vecteur pCEP4 ouvert par les enzymes Not I et Xho I, pour conduire à la construction d'ADN pCEP4-223-RF2.

b) Construction d'ADN codant pour la fusion de la EGFP dans la région cytoplasmique du récepteur α 7-V201-5HT₃.

L'ADNc codant pour la protéine α 7-V201-5HT₃ contient dans son domaine cytoplasmique les séquences des sites Avr II et Pst I respectivement en phase et cohésive avec les séquences clivées par les enzymes Nhe I et Pst I du plasmide pEGFP-C2 (ClonTech) et permettent l'obtention d'un protéine de fusion contenant la séquence de le EGFP dans le domaine cytoplasmique de la protéine α 7-V201-5HT₃.

La construction d'ADN codant pour cette protéine de fusion est donc obtenue par ligation des fragments Not I-Avr II (1036 nt) et Pst I - Xho I (286 nt) de KS 223 avec le fragment Nhe I - Pst I (774 nt) de pEGFP-C2 et le vecteur pCEP4 ouvert par les enzymes Not I et Xho I pour conduire à la construction pCEP4-223-RF3.

Les constructions pCEP4-223-RF1, -RF2, et -RF3 sont ensuite exprimées dans des cellules HEK 293 ainsi que décrit plus haut.

EXEMPLE 9 : CONSTRUCTION D'ADN CODANT POUR UN RÉCEPTEUR DES CHIMIOKINES FLUORESCENT

Le cDNA codant pour le récepteur CCR5 humain des chimiokines (numéro d'accès Genbank : U54994) est amplifié à l'aide des oligonucléotides 5'GGCCCAAGCTTATGTCAGGATCCGGGGAT3' et 5'CGCCCGCTCGAGTCACAAGCCCACAGATAT puis cloné dans le vecteur Bluescript KS ouvert par l'enzyme de restriction Eco RV.

L'insert est excisé par les enzymes Hind III ou BamH1 (5') et Xho I (3') (fragment 1) pour être cloné en phase avec le cDNA codant pour la EGFP ou la ECFP.

La fusion en phase de la EGFP ou la ECFP avec le peptide signal du plasmide pJL223 (exemple 1) et le gène codant pour le récepteur CCR5, est effectuée de la manière suivante. Le plasmide pJL223 est ouvert par l'enzyme de restriction Age I puis rendu franc par mélange avec la Klenow et des nucléotides. Le fragment s'étendant jusqu'au site Xho I est ensuite éliminé (coupe Xho I) (fragment 2).

Les plasmides pEGFP-C3 et pECFP-C3 sont ouverts par l'enzyme Age I et les extrémités sont rendues franches. Les inserts Age I franc- Hind III et Age I franc- Xho I (fragments 3) sont ligés avec les fragments 1 et 2 pour donner

lieu au constructions KS-SP EGFP-Hind III-CCR5, KS-SP EGFP-Xho I-CCR5, KS-SP ECFP-Hind III-CCR5 et KS-SP ECFP-Xho I-CCR5. Les ADNs correspondant aux séquences SP EGFP-Hind III-CCR5, SP EGFP-Xho I-CCR5, SP ECFP-Hind III-CCR5, et SP ECFP-Xho I-CCR5 sont excisés par Not I et Xho I puis clonés dans le vecteur pCEP4 ouvert pas les mêmes enzymes.

La figure 16 montre le spectre d'émission de cellules HEK 293 transfectées et exprimant le récepteur EGFP-Hind III-CCR5, et la figure 17 montre la réponse de libération de calcium intracellulaire de cellules exprimant la construction d'AND EGFP-HindCCR5.

EXEMPLE 10 : CONSTRUCTION D'ADN CODANT POUR LA CHIMIOKINE HUMAINE RANTES FLUORESCENTE

- 15 L'ADNc codant pour la chimiokine RANTES humaine (numéro d'accès Genbank M21121) est synthétisé par PCR récursive à l'aide des oligonucléotides
- 1 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90
- 1 5'GTTGACAAGCTTCGGGATCCA3'
 - 2 5'AGCACAGAGGGCAGTAGCAATGAGGATGACAGCG
AGGCGTGCCGCCGGAGACCTTCATTGGATCCGAAGCTTGTCAAC3'
 - 3 5'ATTGCTACTGCCCTCTGTGCTCCTGCATCTGCCTCC
CCATATTCCCTCGGACACCACACCATGCTGCTTCGCCTACATT3'
 - 4 5'GCACTTGCCACTGGTGTAGAAATACTCCTGATGTGG
GCACGGGGCAGTGGCGGGCAATGTAGGCGAACGCAGCATGG3'
 - 5 5'GCACTTGCCACTGGTGTAGAAATACTCCTGATGT
GGGCACGGGGCAGTGGCGGGCAATGTAGGCGAACGCAGCATGG3'
 - 6 5'CTAGCTCATCTCCAGCGAGTTGATGTACTCCCGAACCC
CATTCTTCTCTGGGTTGGCACAAACTTGACG3'
 - 7 5'AACTCGCTGGAGATGAGCTAGGCGGCCGCTCG
AGGTCGACCTAGTCACTA3'

8 5'TAGTGACTAGGTCGACCTCGA3'

selon les protocoles décrits dans Prodromou & Pearl, Protein Engineering Vol 5, pp 827-829, 1992, puis cloné dans le vecteur Bluescript KS ouvert par Eco RV.

Les ADN codant respectivement pour la chimiokine RANTES (fragment Hind III - Not I ou Bam HI-Not I) et pour la protéine fluorescente jaune (fragments Age I franc - Bam HI ou Age I franc - Hin DIII du plasmide pEYFP-C3) sont clonés en phase avec le peptide signal de l'alpha-mating factor de levure entre les sites Sna BI et Not I du vecteur pPIC 9, pour conduire aux constructions pPIC9-EYFP- HindIII-RANTES et pPIC9-EYFP-Bam HI-RANTES.

La figure 18 montre un immunoblot révélant l'expression de la construction EYFP- HindIII-RANTES dans un surnageant de culture de levure *pichia pastoris* transformée par le plasmide pPIC9-EYFP- HindIII-RANTES, et détectée à l'aide d'une anticorps anti-GFP.

**EXEMPLE 11 : IDENTIFICATION DE NOUVEAUX LIGANDS POUR
DES RECEPTEURS ORPHELINS**

1) fusion de la EGFP avec l'extrémité carboxyterminale du récepteur NK2R des tachykinines.

l'ADNc codant pour le récepteur NK2R compris dans le plasmide pKS NK2R (exemple 1) est clivé par l'enzyme Age I. L'extrémité 5' dépassante obtenue est rendu franche à l'aide de nucleotides et de la Klenow polymerase. L'ADN linéarisé est clivé par l'enzyme Not I pour produire un fragment de 1045 nt (fragment a).

Le plasmide pEGFP-C3 est ouvert à l'aide de l'enzyme Age I puis rendu franc. Il est ensuite clivé par l'enzyme Xho I pour conduire à un fragment b, de 739 nt.

Les fragments a et b sont ligués avec le vecteur pCEP4 ouvert par les enzymes Not I et Xho I pour conduire au plasmide pCEP4-NK2R-AgeI-EGFP.

35

2) fusion de la protéine fluorescente EYFP avec l'extrémité aminoterminale de la sous-unité alpha de la protéine Gq de souris.

L'ADNc codant pour la sous-unité alpha de la protéine Gq de souris (Genbank accession N° M55412) est amplifiée par PCR à l'aide des l'oligonucléotides

5'GC GGCCGCATGGGGATCCTACTCTGGAGTCCATGGCG et 5'
5'CCGCTCGAGTTAATCTAGAAGGACCAGATTGTACTCCTTCAGG afin
d'introduire des sites Not I et Bam HI en 5' et des sites Xho I et Xba I en 3' du
gène codant pour la sous-unité alpha de la protéine Gq.

Le produit PCR obtenu est cloné dans le vecteur KS ouvert par Eco RV.

Le cDNA codant pour la sous-unité alpha Gq est excisé par les enzymes
BamHI et Xba I et ligué dans le plasmide pEYFP-C3 ouvert par les enzymes
Bgl II et Xba I pour conduire au plasmide pEYFP-Bgl II-Gq .

3) coexpression du récepteur NK2R-Age I-EGFP et de la protéine EYFP-Bgl II-Gq .

15 Des cellules HEK 293 sont transfectées à l'aide des plasmides pEYFP-Bgl II-Gq ou pCEP4-NK2R-AgeI-EGFP et pEYFP-Bgl II-Gq . Les cellules exprimant le produit des gènes du plasmide pCEP4 sont sélectionnées à l'aide de l'antibiotique hygromycine B, celles qui expriment les produits du gène du
20 plasmide pEGFP modifié sont sélectionnées à l'aide de néomycine.

4) tests de fonctionnalité de la protéine Gq fluorescente.

Ils sont effectués par le biais d'études fonctionnelles de libération de calcium. La surexpression de protéines G entraîne une augmentation de
25 l'amplitude des réponses calciques.

30 5) les cellules exprimant la protéine Gq de manière stable sont transfectées à l'aide d'un plasmide d'expression codant pour un récepteur orphelin. L'effet de molécules est mesuré par fluorescence à 405 nm en excitant la suspension de cellules à 355 nm.

RE VENDICATIONS

1. Utilisation d'une protéine fluorescente choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est supérieur à environ $14000 M^{-1} cm^{-1}$ et le rendement quantique de fluorescence est supérieur à environ 0,38, notamment choisie parmi :

- la protéine fluorescente verte (GFP), ou
- des variants dérivés de la GFP, par addition, délétion ou substitution d'un ou plusieurs acides aminés, sous réserve que ces variants conservent la propriété de fluorescence,

pour la détection et la quantification d'interactions non covalentes entre une protéine cible marquée par voie génétique par la GFP ou l'un des variants ci-dessus définis ou l'un des fragments ci-dessus définis et l'un de ses ligands marqué par un marqueur constitué :

- soit par une molécule susceptible d'absorber la lumière émise par la protéine fluorescente,

20 cette détection et quantification ayant lieu par transfert d'énergie de fluorescence:

- . entre la GFP ou l'un des variants définis ci-dessus, ou l'un des fragments définis ci-dessus et la susdite substance fluorescente, la substance fluorescente étant telle que soit elle est excitable à la longueur d'onde d'émission de la GFP ou de l'un des susdits variants, ou de l'un des susdits fragments, soit elle émet à la longueur d'excitation de la GFP, ou de l'un des susdits variants, ou de l'un des susdits fragments, ou

- . entre la GFP ou l'un de ses variants définis ci-dessus, ou l'un des fragments définis ci-dessus, et la susdite molécule susceptible d'absorber la lumière émise par la protéine fluorescente.

35 2. Utilisation d'un ligand marqué par un marqueur constitué :

- soit par une molécule susceptible d'absorber la lumière émise par la protéine fluorescente,
- soit par une substance fluorescente,

pour la détection et la quantification d'interactions non covalentes entre une protéine cible et le susdit ligand, ladite protéine cible étant marquée par voie génétique par une protéine fluorescente choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est supérieur à environ $14000\text{M}^{-1}\text{cm}^{-1}$ et le rendement quantique de fluorescence est supérieur à environ 0,38, notamment choisie parmi :

- la protéine fluorescente verte (GFP), ou
- des variants dérivés de la GFP, par addition, délétion ou substitution d'un ou plusieurs acides aminés, sous réserve que ces variants conservent la propriété de fluorescence,
- ou des fragments de la GFP, ou de fragments des susdits variants, sous réserve que ces fragments conservent la propriété de fluorescence, cette détection et quantification ayant lieu par transfert d'énergie de fluorescence:
 - . entre la GFP ou l'un des variants définis ci-dessus, ou l'un des fragments définis ci-dessus et la susdite substance fluorescente, la substance fluorescente étant telle que soit elle est excitable à la longueur d'onde d'émission de la GFP ou de l'un des susdits variants, ou de l'un des susdits fragments, soit elle émet à la longueur d'excitation de la GFP, ou de l'un des susdits variants, ou de l'un des susdits fragments, ou
 - . entre la GFP ou l'un de ses variants définis ci-dessus, ou l'un des fragments définis ci-dessus, et la susdite molécule susceptible d'absorber la lumière émise par la protéine fluorescente.

3. Utilisation selon la revendication 1, dans laquelle la protéine fluorescente est choisie parmi :

- la protéine fluorescente verte (GFP ou EGFP),
 - la protéine fluorescente cyan (CFP ou ECFP),
 - la protéine fluorescente jaune (YFP ou EYFP),
 - GFPUV,
- ou leurs mutants dans lesquels les codons sont optimisés pour l'expression dans les cellules humaines, bactériennes ou végétales,
- ou leurs mutants présentant des longueurs d'excitation ou d'émission plus élevées ou plus faibles que celles associées aux protéines définies ci-dessus, sous réserve que leur coefficient d'extinction moléculaire soit supérieur à environ

14000M⁻¹cm⁻¹ et leur rendement quantique de fluorescence soit supérieur à environ 0,38.

4. Utilisation d'une protéine fluorescente (n° 1) selon l'une des revendications 1 ou 3, dans laquelle le ligand est marqué

* soit par une substance fluorescente, le marquage étant :

- soit effectué par voie chimique, la substance fluorescente étant alors un composé chimique,

- soit effectué par voie recombinante, la substance fluorescente étant alors un peptide ou une protéine fluorescente (n° 2), et pouvant être notamment choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est supérieur à environ 14000M⁻¹cm⁻¹ et le rendement quantique de fluorescence est supérieur à environ 0,38, notamment choisie parmi :

- la protéine fluorescente verte (GFP), ou

- des variants dérivés de la GFP, par addition, délétion ou substitution d'un ou plusieurs acides aminés, sous réserve que ces variants conservent la propriété de fluorescence,

- ou des fragments de la GFP, ou de fragments des susdits variants, sous réserve que ces fragments conservent la propriété de fluorescence,

* soit par une substance non fluorescente appartenant au groupe des acides violets [Acid Violet 5, CAS 10130-48-0 ; Acid Violet 7, CAS 4321-69-1 ; Acid Violet 17, CAS 4129-84-4], acides rouges [Acid Red 1, CAS 3734-67-6 ; Acid Red 8, CAS 4787-93-3 ; Acid Red 37, CAS 6360-07-2 ; Acid Red 40, CAS 12167-45-2 ; Acid Red 106, CAS 6844-74-2 ; Acid Red 114, CAS 6459-94-5], les alizarines, l'aluminon, l'azocarmine B [CAS 25360-72-9], la fuschine basique [Basic Red 9, CAS 569-61-9], le Bordeaux R [Acid Red 17, CAS 5858-33-3], la Carmine [CAS 1390-65-4].

5. Utilisation d'une protéine fluorescente selon l'une des revendications 1, 3 ou 4 dans laquelle la protéine cible et le ligand sont marqués par voie génétique, la protéine fluorescente et la substance fluorescente étant respectivement choisies parmi les couples de composés suivants :

GFPUV - EYFP

EYFP - GFPUV

ECFP - EYFP

EYFP - ECFP

ECFP - EGFP

EGFP - ECFP

EGFP - EYFP

EYFP - EGFP

et notamment dans laquelle la protéine cible est marquée par la protéine EYFP ou EGFP et le ligand est marqué par la protéine ECFP, ou la protéine cible est marquée par la protéine ECFP et le ligand est marqué par la protéine EYFP ou EGFP.

6. Utilisation selon la revendication 1 d'une protéine fluorescente choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est supérieur à environ $14000\text{M}^{-1}\text{cm}^{-1}$ et le rendement quantique de fluorescence est supérieur à environ 0,38, notamment choisie parmi :

- la protéine fluorescente verte (GFP), ou
- des variants dérivés de la GFP, par addition, délétion ou substitution d'un ou plusieurs acides aminés, sous réserve que ces variants conservent la propriété de fluorescence,
- ou des fragments de la GFP, ou de fragments des susdits variants, sous réserve que ces fragments conservent la propriété de fluorescence,

pour la détection et la quantification d'interactions non covalentes entre une protéine cible marquée par voie génétique par la GFP ou l'un des variants ci-dessus définis ou l'un des fragments ci-dessus définis et l'un de ses ligands marqué par une substance fluorescente, cette détection et quantification ayant lieu par transfert d'énergie de fluorescence entre la GFP ou l'un des variants définis ci-dessus, ou l'un des fragments définis ci-dessus et la susdite substance fluorescente, la substance fluorescente étant telle que soit elle est excitable à la longueur d'onde d'émission de la GFP ou de l'un des susdits variants, ou de l'un des susdits fragments, soit elle émet à la longueur d'onde d'excitation de la GFP, ou de l'un des susdits variants, ou de l'un des susdits fragments.

30

7. Utilisation selon l'une des revendications 1 à 6, dans laquelle la protéine fluorescente est EGFP et dans laquelle :

- soit la EGFP est donneur d'énergie de fluorescence et le marqueur absorbant la lumière émise par la EGFP est une substance fluorescente ou non, et le marqueur étant choisi parmi des substances, dont le spectre d'excitation chevauche le spectre d'émission de la EGFP, et notamment dans le cas où le marqueur est une substance fluorescente, il est choisi parmi : le 4,4-difluoro-4-bora-3a,4a-diaza-s-indacene (Bodipy), l'éosine, l'érythrosine, la

tétraméthylrhodamine, la sulforhodamine 101 commercialisée par Molecular probe sous la dénomination Texas Red, et leurs dérivés permettant d'une part le greffage et, d'autre part, dont le spectre d'excitation recouvre le spectre d'émission de EGFP,

5

10

15

20

25

30

et dans le cas où le marqueur n'est pas une substance fluorescente, il est choisi parmi le groupe des acides violets [Acid Violet 5, CAS 10130-48-0 ; Acid Violet 7, CAS 4321-69-1 ; Acid Violet 17, CAS 4129-84-4], acides rouges [Acid Red 1, CAS 3734-67-6 ; Acid Red 8, CAS 4787-93-3 ; Acid Red 37, CAS 6360-07-2 ; Acid Red 40, CAS 12167-45-2 ; Acid Red 106, CAS 6844-74-2 ; Acid Red 114, CAS 6459-94-5], les alizarines, l'aluminon, l'azocarmine B [CAS 25360-72-9], la fuschine basique [Basic Red 9, CAS 569-61-9], le Bordeaux R [Acid Red 17, CAS 5858-33-3], la Carmine [CAS 1390-65-4],

- soit la EGFP est accepteur d'énergie de fluorescence et la substance fluorescente est donneur d'énergie de fluorescence et est choisie parmi des substances, dont le spectre d'émission chevauche le spectre d'excitation de la EGFP, et notamment parmi : les coumarines, la fluorescamine, le 6-(N-méthylanilino)naphtalène, (mansyl) et leurs dérivés permettent d'une part le greffage et, d'autre part, dont le spectre d'émission recouvre le spectre d'excitation de EGFP,

- ou soit la protéine fluorescente est ECFP et est donneur d'énergie de fluorescence et la substance fluorescente est accepteur d'énergie et est choisi parmi la fluorescéine et le 7-nitrobenz-2-oxa-1,3-diazole,

- ou soit la protéine fluorescente est ECFP et est accepteur d'énergie de fluorescence et la substance fluorescente est donneur d'énergie et est choisie parmi le pyrène ou la coumarine ou leurs dérivés permettant d'une part le greffage, et, d'autre part, dont le spectre d'émission chevauche le spectre d'excitation de la ECFP.

8. Utilisation selon l'une des revendications 1 à 7, dans laquelle la protéine cible est choisie parmi :

- les récepteurs membranaires couplés à la protéine G, notamment dans Supplement Trends in Pharmacological Sciences, 1997 (*Receptor and ion Channel Nomenclature*),

- les récepteurs des facteurs de croissance, notamment ceux qui sont structurellement reliés au récepteur de l'insuline (Yarden, Y. and Ullrich, A. 1988, Biochemistry 27:3113-3119) ou au récepteur de l'interféron γ (Brisco,

J. et al. 1996, *Phylos. Trans. R. Soc. Lond. B. Biol. Sci.* 351:167-171 ; Ihle, J.N. 1995, *Nature* 377:591-594),

- les récepteurs canaux, notamment dans *Supplement Trends in Pharmacological Sciences*, 1997 (*Receptor and ion Channel Nomenclature*),

5 - les récepteurs nucléaires intracellulaires, notamment ceux qui sont structurellement reliés au récepteur des stéroïdes (Mangelsdorf et al. 1995, *Cell*, 83:835-839 ; Wurtz, J.L. et al. 1996, *Nature Struct. Biol.* 3:206).

10 9. Utilisation selon l'une des revendications 1 à 8, dans laquelle la protéine cible est choisie parmi les récepteurs membranaires couplés à la protéine G.

15 10. Procédé de détection et de quantification d'interactions non covalentes entre une protéine cible, notamment un récepteur, et l'un de ses ligands, caractérisé en ce que :

20 - on prépare des cellules soit des fragments de cellules exprimant une séquence d'ADN comprenant le gène codant pour une protéine fluorescente fusionnée avec le gène de la protéine cible, la fusion entre le gène de la protéine fluorescente et le gène de la susdite protéine cible étant telle que les propriétés de la protéine cible, notamment du récepteur, ne sont pas modifiées par la présence de la protéine fluorescente, à savoir :

* l'interaction entre la protéine cible, notamment le récepteur, et le ligand n'est pas modifiée,

25 * la fonction de transduction de la réponse n'est pas modifiée, la protéine fluorescente étant choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est supérieur à environ 14000M⁻¹cm⁻¹ et le rendement quantique de fluorescence est supérieur à environ 0,38, notamment choisie parmi :

30 - la protéine fluorescente verte (GFP), ou

- des variants dérivés de la GFP, par addition, délétion ou substitution d'un ou plusieurs acides aminés, sous réserve que ces variants conservent la propriété de fluorescence,

35 - ou des fragments de la GFP, ou de fragments des susdits variants, sous réserve que ces fragments conservent la propriété de fluorescence,

- on met en présence les susdites cellules ou les susdits fragments de cellules avec un ligand de la susdite protéine cible, notamment du susdit récepteur, marqué par un marqueur constitué :

- soit par une molécule susceptible d'absorber la lumière émise par la protéine fluorescente,

- soit par une substance fluorescente,

et soit la protéine fluorescente étant donneur d'énergie de fluorescence et le marqueur étant accepteur d'énergie de fluorescence, ou soit la protéine fluorescente étant accepteur d'énergie de fluorescence et le marqueur étant une substance fluorescente donneur d'énergie de fluorescence, et

- on irradie à une longueur d'onde permettant soit d'exciter la protéine fluorescente, soit d'exciter la substance fluorescente,

- les susdites étapes de mise en présence et d'irradiation pouvant être effectuées soit simultanément, soit l'une après l'autre, ou

- on met en présence les susdites cellules ou les susdits fragments de cellules avec un ligand de la susdite protéine, notamment du susdit récepteur, marqué par un marqueur, les cellules ou le ligand ayant été irradiés préalablement à leur mise en présence,

- on détecte soit une diminution d'amplitude de l'émission du donneur et/ou un signal d'émission caractéristique de l'émission de l'accepteur.

20 11. Procédé de détection et de quantification d'interactions non covalentes entre une protéine cible, notamment un récepteur, et l'un de ses ligands, caractérisé en ce que :

- on prépare une protéine fluorescente fusionnée avec une protéine cible dont on veut déterminer l'interaction protéine-ligand, la fusion entre la protéine fluorescente et la susdite protéine cible étant telle que les propriétés de la protéine, notamment du récepteur, ne sont pas modifiées par la présence de la protéine fluorescente, à savoir :

* l'interaction entre la protéine cible, notamment le récepteur, et le ligand n'est pas modifiée,

* la fonction de transduction de la réponse n'est pas modifiée,

la protéine fluorescente étant choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est supérieur à environ $14000\text{M}^{-1}\text{cm}^{-1}$ et le rendement quantique de fluorescence est supérieur à environ 0,38, notamment choisie parmi :

- la protéine fluorescente verte (GFP), ou

- des variants dérivés de la GFP, par addition, délétion ou substitution d'un ou plusieurs acides aminés, sous réserve que ces variants conservent la propriété de fluorescence,

5 - ou des fragments de la GFP, ou de fragments des susdits variants, sous réserve que ces fragments conservent la propriété de fluorescence,

- on met en présence la susdite protéine fluorescente fusionnée avec la protéine cible avec un ligand de la susdite protéine, notamment du susdit récepteur, ce ligand étant marqué par un marqueur constitué :

10 - soit par une molécule susceptible d'absorber la lumière émise par la protéine fluorescente,

- soit par une substance fluorescente,

et soit la protéine fluorescente étant donneur d'énergie de fluorescence et le marqueur étant accepteur d'énergie de fluorescence; ou soit la protéine fluorescente étant accepteur d'énergie de fluorescence et le marqueur est une substance fluorescente donneur d'énergie de fluorescence, et

15 - on irradie à une longueur d'onde permettant soit d'exciter la protéine fluorescente, soit d'exciter la substance fluorescente,

- les susdites étapes de mise en présence et d'irradiation pouvant être effectuées soit simultanément, soit l'une après l'autre, ou

20 - on met en présence la susdite protéine fluorescente fusionnée avec la protéine cible avec un ligand de la susdite protéine, notamment du susdit récepteur, ce ligand étant marqué par un marqueur constitué :

- soit par une molécule susceptible d'absorber la lumière émise par la protéine fluorescente,

25 - soit par une substance fluorescente,

la protéine fluorescente fusionnée avec la protéine cible ou le ligand ayant été irradiés préalablement à leur mise en présence,

- on détecte soit une diminution d'amplitude de l'émission du donneur et/ou un signal d'émission caractéristique d'émission de l'accepteur.

30

12. Procédé selon l'une des revendication 10 ou 11, dans lequel on introduit une étape supplémentaire :

- avant ou après, ou simultanément à l'étape de mise en présence de la protéine fluorescente fusionnée avec la protéine cible et d'un ligand de la susdite protéine, ce ligand étant marqué par un marqueur, ou

35 - avant, ou après, ou simultanément à l'étape de mise en présence de cellule ou de fragments de cellules et d'un ligand de la susdite protéine, marqué par un marqueur,

cette étape supplémentaire consistant :

- soit à mettre la susdite protéine fluorescente fusionnée avec la protéine cible en présence du susdit ligand non marqué et simultanément du susdit ligand marqué,
- 5 - soit à mettre en présence les susdites cellules ou les susdits fragments de cellules en présence simultanée du susdit ligand non marqué et du susdit ligand marqué,
- on détecte soit une diminution d'amplitude de l'émission du donneur et/ou un signal d'émission caractéristique de l'émission de l'accepteur, 10 respectivement dans le cas de l'utilisation du ligand marqué et dans le cas de l'utilisation simultanée du ligand marqué et du ligand non marqué,
- et on compare soit les diminutions d'amplitude de l'émission du donneur respectivement obtenues et/ou les signaux d'émission caractéristiques de l'émission de l'accepteur respectivement obtenus.

15

13. Procédé selon la revendication 12, dans lequel la protéine fluorescente est EGFP et dans lequel :

- soit la EGFP est donneur d'énergie de fluorescence et le marqueur est accepteur d'énergie de fluorescence et est choisie parmi des substances, dont le spectre d'excitation chevauche le spectre d'émission de la EGFP, et notamment dans le cas où le marqueur est une substance fluorescente, il est choisi parmi : le 4,4-difluoro-4-bora-3a,4a-diaza-s-indacene (Bodipy), l'éosine, l'érythrosine, la tétraméthylrhodamine, la sulforhodamine 101 commercialisée par Molecular probe sous la dénomination Texas Red, et leurs dérivés permettent d'une part le greffage et, d'autre part, dont le spectre d'excitation recouvre le spectre d'émission de EGFP,

20

et dans le cas où le marqueur n'est pas une substance fluorescente, il est choisi parmi le groupe des acides violets [Acid Violet 5, CAS 10130-48-0 ; Acid Violet 7, CAS 4321-69-1 ; Acid Violet 17, CAS 4129-84-4], acides rouges [Acid Red 1, CAS 3734-67-6 ; Acid Red 8, CAS 4787-93-3 ; Acid Red 37, CAS 6360-07-2 ; Acid Red 40, CAS 12167-45-2 ; Acid Red 106, CAS 6844-74-2 ; Acid Red 114, CAS 6459-94-5], les alizarines, l'aluminon, l'azocarmine B [CAS 25360-72-9], la fuschine basique [Basic Red 9, CAS 569-61-9], le Bordeaux R [Acid Red 17, CAS 5858-33-3], la Carmine [CAS 1390-65-4],

35

- soit la EGFP est accepteur d'énergie de fluorescence et la substance fluorescente est donneur d'énergie de fluorescence et est choisie parmi

des substances, dont le spectre d'émission chevauche le spectre d'excitation de la EGFP, et notamment parmi : les coumarines, la fluorescamine, le 6-(N-méthylanilino)naphtalène, (mansyl) et leurs dérivés permettent d'une part le greffage et, d'autre part, dont le spectre d'émission recouvre le spectre d'excitation de EGFP.

14. Procédé selon l'une des revendications 10 à 13, dans lequel la protéine dont on veut déterminer l'interaction protéine-ligand est choisie parmi:

10 - les récepteurs membranaires couplés à la protéine G, notamment dans Supplement Trends in Pharmacological Sciences, 1997 (*Receptor and ion Channel Nomenclature*),

15 - les récepteurs des facteurs de croissance, notamment ceux qui sont structurellement reliés au récepteur de l'insuline (Yarden, Y. and Ullrich, A. 1988, Biochemistry 27:3113-3119) ou au récepteur de l'interféron γ (Brisco, J. et al. 1996, Phylos. Trans. R. Soc. Lond. B. Biol. Sci. 351:167-171 ; Ihle, J.N. 1995, Nature 377:591-594),

20 - les récepteurs canaux, notamment dans Supplement Trends in Pharmacological Sciences, 1997 (*Receptor and ion Channel Nomenclature*),

25 - les récepteurs nucléaires intracellulaires, notamment ceux qui sont structurellement reliés au récepteur des stéroïdes (Mangelsdorf et al. 1995, Cell, 83:835-839 ; Wurtz, J.L. et al. 1996, Nature Struct. Biol. 3:206).

15. Procédé selon l'une des revendications 10 à 14, dans lequel la protéine fluorescente est la EGFP et la substance marquée est le Bodipy et dans lequel on détecte soit la diminution d'amplitude d'émission de la EGFP, soit le signal d'émission du Bodipy résultant du transfert d'énergie, la longueur d'onde d'irradiation correspondant à la longueur d'onde d'excitation de la EGFP.

30 16. Procédé selon l'une des revendications 10 à 15, dans lequel la protéine fluorescente est la EGFP et la substance marquée est une coumarine et dans lequel on détecte soit la diminution d'amplitude de la coumarine, soit le signal d'émission de la EGFP résultant du transfert d'énergie, la longueur d'onde d'irradiation correspondant à la longueur d'onde d'excitation de la coumarine.

17. Procédé selon l'une des revendications 10 à 16, dans lequel la protéine fluorescente est fusionnée du côté N-terminal et la protéine cible, notamment le récepteur, est fusionné du côté C-terminal.

5 18. Procédé selon l'une des revendications 10 à 17, dans lequel la protéine fluorescente est fusionnée du côté C-terminal et la protéine cible, notamment le récepteur, est fusionné du côté N-terminal.

10 19. Procédé selon l'une des revendications 10 à 18, dans lequel la protéine fluorescente est insérée dans la protéine cible à un endroit ne correspondant pas à un site de liaison protéine cible-ligand, notamment dans le cas des récepteurs couplés à la protéine G, cette insertion ayant lieu dans la première ou la troisième boucle intracellulaire du récepteur, sous réserve que l'insertion ne détruise ni les propriétés du récepteur, ni la fluorescence de la protéine fluorescente.

20 20. Procédé selon l'une des revendications 10 à 19, dans lequel les cellules sont des cellules de mammifères, notamment les cellules HEK 293 adhérentes ou en suspension, cellules CHO, cellules COS, lignées lymphocytaires, fibroblastes, etc., ou des cellules de levure, notamment *pichia* telle que *pichia pastoris*, *saccharomyces* telle que *saccharomyces cerevisiae*, *saccharomyces kluyveri*, *Hansenula* telle que *hansenula polymorpha*, ou des cellules d'insectes infectées par un virus tel que *baculovirus*, notamment cellules TNI ou sf9, ou des champignons, notamment les souches de *Aspergillus* (*A. oryzae*, *A. nidulans*, *A. niger*), *Neurospora*, *Fusarium*, *Trichoderma*.

30 21. Procédé selon l'une quelconque des revendications 10 à 20, dans lequel un signal est détectable, dans un appareil de fluorimétrie conventionnel ou dans un appareil de mélange rapide équipé d'un système de détection de fluorescence, après mélange du donneur et de l'accepteur et peut être aboli par l'addition d'une substance non fluorescente de même spécificité pharmacologique, et notamment dans lequel le rapport signal/bruit est supérieur à environ 2.

35 22. Utilisation d'un protéine fluorescente choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est supérieur à environ

14000 M⁻¹ cm⁻¹ et le rendement quantique de fluorescence est supérieur à environ 0,38, notamment choisie parmi :

- la protéine fluorescente verte (GFP), ou
 - des variants dérivés de la GFP, par addition, délétion ou remplacement d'un ou plusieurs acides aminés, sous réserve que ces variants conservent la propriété de fluorescence.

- ou des fragments de la GFP, ou de fragments des susdits variants, sous réserve que ces fragments conservent la propriété de fluorescence, pour la détection et la quantification d'interactions non covalentes entre une protéine cible constituée par un récepteur couplé aux protéines G et une protéine G, en vue d'identifier les molécules biologiquement actives vis-à-vis du récepteur, susceptibles de former une interaction non covalente réversible avec ledit récepteur, ledit récepteur étant marqué par voie génétique par la protéine fluorescente et la protéine G étant marquée par un marqueur constitué :

- soit par une molécule susceptible d'absorber la lumière émise par protéine fluorescente,

- soit par une substance fluorescente pouvant être notamment choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est supérieur à environ $14000M^{-1}cm^{-1}$ et le rendement quantique de fluorescence est supérieur à environ 0,38, notamment choisie parmi :

- la protéine fluorescente verte (GFP), ou
 - des variants dérivés de la GFP, par addition , délétion ou
d'un ou plusieurs acides aminés, sous réserve que ces variants
la propriété de fluorescence.

- ou des fragments de la GFP, ou de fragments des susdits variants, sous réserve que ces fragments conservent la propriété de fluorescence.

cette détection et quantification ayant lieu par transfert d'énergie de fluorescence entre le récepteur marqué par la GFP ou l'un de ses variants définis ci-dessus, ou l'un des fragments définis ci-dessus et la susdite substance fluorescente, la substance fluorescente étant telle que soit elle est excitable à la longueur d'onde d'émission de la GFP ou de l'un des susdits variants, ou de l'un des susdits fragments, soit elle émet à la longueur d'excitation de la GFP, ou de l'un des susdits variants, ou de l'un des susdits fragments, ou

- entre la GFP ou l'un de ses variants définis ci-dessus, ou l'un des fragments définis ci-dessus, et la susdite molécule susceptible d'absorber la lumière émise par la protéine fluorescente.

23. Utilisation d'une protéine G marquée par un marqueur constitué:

- soit par une molécule susceptible d'absorber la lumière émise par la protéine fluorescente,

5 - soit par une substance fluorescente pouvant être notamment choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est supérieur à environ $14000\text{M}^{-1}\text{cm}^{-1}$ et le rendement quantique de fluorescence est supérieur à environ 0,38, notamment choisie parmi :

10 - la protéine fluorescente verte (GFP), ou

- des variants dérivés de la GFP, par addition, délétion ou substitution d'un ou plusieurs acides aminés, sous réserve que ces variants conservent la propriété de fluorescence,

15 - ou des fragments de la GFP, ou de fragments des susdits variants, sous réserve que ces fragments conservent la propriété de fluorescence,

pour la détection et la quantification d'interactions non covalentes entre une protéine cible constituée par un récepteur couplé aux protéines G et la susdite protéine G, en vue d'identifier les molécules biologiquement actives vis-à-vis du récepteur, susceptibles de former une interaction non covalente réversible avec ledit récepteur, ledit récepteur étant marqué par voie génétique par une protéine fluorescente choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires dont le coefficient d'extinction moléculaire est supérieur à environ $14000\text{M}^{-1}\text{cm}^{-1}$ et le rendement quantique de fluorescence est supérieur à environ 0,38, notamment choisie parmi :

20 - la protéine fluorescente verte (GFP), ou

- des variants dérivés de la GFP, par addition, délétion ou substitution d'un ou plusieurs acides aminés, sous réserve que ces variants conservent la propriété de fluorescence,

25 - ou des fragments de la GFP, ou de fragments des susdits variants, sous réserve que ces fragments conservent la propriété de fluorescence,

30 cette détection et quantification ayant lieu par transfert d'énergie de fluorescence:

35 . entre la GFP ou l'un des variants définis ci-dessus, ou l'un des fragments définis ci-dessus et la susdite substance fluorescente, la substance fluorescente étant telle que soit elle est excitable à la longueur d'onde d'émission de la GFP ou de l'un des susdits variants, ou de l'un des susdits fragments, soit elle émet à la

longueur d'excitation de la GFP, ou de l'un des susdits variants, ou de l'un des susdits fragments, ou

entre la GFP ou l'un de ses variants définis ci-dessus, ou l'un des fragments définis ci-dessus, et la susdite molécule susceptible d'absorber la lumière émise par la protéine fluorescente.

5

10

24. Utilisation selon la revendication 22 ou 23, dans laquelle la protéine fluorescente est choisie parmi EGFP, ECFP, EYFP, ou leurs mutants pour lesquels le coefficient d'extinction moléculaire est supérieur à environ 14000 M⁻¹cm⁻¹ et le rendement quantique est supérieur à environ 0,38.

25. Utilisation selon l'une des revendications 22 ou 24, dans laquelle le récepteur est choisi parmi :

15

- les récepteurs couplés aux protéines G, notamment ceux décrits dans Supplément Trends in Pharmacological Sciences, 1997 (*Receptor and ion Channel Nomenclature*),

20

- les séquences codant pour des récepteurs couplés aux protéines G putatifs dont les molécules biologiquement actives vis-à-vis d'eux-mêmes sont à identifier, notamment choisies parmi les séquences de récepteur orphelins disponibles dans les banques de séquence Genbank, EMBL et banques affiliées.

25

26. Utilisation selon la revendication 23, dans laquelle la protéine G est choisie parmi les protéines G décrites dans Journal of Receptor Research, vol 13, pp19-26, 1993 ou Angewandte Chemie, ed. Engl. vol 34, pp1406-1419, 1995.

30

27. Procédé de détection et de quantification d'interactions non covalentes entre une protéine cible constituée par un récepteur couplé aux protéines G et une protéine G, en vue d'identifier les molécules biologiquement actives vis-à-vis du récepteur, susceptibles de former une interaction non covalente réversible avec ledit récepteur, caractérisé en ce que :

35

- on prépare des cellules soit des fragments de cellules exprimant une séquence d'ADN comprenant le gène codant pour une protéine fluorescente fusionnée avec le gène du récepteur couplé aux protéines G, la fusion entre le gène de la protéine fluorescente et le gène du susdit récepteur étant telle que les propriétés du récepteur ne sont pas modifiées par la présence de la protéine fluorescente, à savoir :

* l'interaction entre le récepteur et la protéine G n'est pas modifiée,

* l'interaction entre le récepteur et la molécule biologiquement active n'est pas modifiée,

* la fonction de transduction de la réponse n'est pas modifiée, la protéine fluorescente étant choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires dont le coefficient d'extinction moléculaire est supérieur à environ $14000 M^{-1} cm^{-1}$ et le rendement quantique de fluorescence est supérieur à environ 0,38, notamment choisie parmi :

- la protéine fluorescente verte (GFP), ou

- des variants dérivés de la GFP, par addition, délétion ou substitution d'un ou plusieurs acides aminés, sous réserve que ces variants conservent la propriété de fluorescence,

- ou des fragments de la GFP, ou de fragments des susdits variants, sous réserve que ces fragments conservent la propriété de fluorescence,

la protéine G étant marquée par un marqueur constitué :

- soit par une molécule susceptible d'absorber la lumière émise par la protéine fluorescente,

- soit par une substance fluorescente,

- la protéine fluorescente et le susdit marqueur étant tels qu'ils peuvent transférer l'énergie de l'un à l'autre, la protéine fluorescente pouvant être donneur d'énergie ou le susdit marqueur pouvant être donneur d'énergie, l'interaction entre le récepteur marqué par la protéine fluorescente et la protéine G marquée par un marqueur susdéfini étant détectée par transfert d'énergie de fluorescence.

25

28. Procédé d'identification et éventuellement de quantification d'interactions entre un récepteur et une molécule non fluorescente biologiquement active vis-à-vis dudit récepteur, susceptibles de former une interaction non covalente réversible avec ledit récepteur, par mise en œuvre du procédé selon la revendication 27, dans lequel une molécule non fluorescente biologiquement active est ajoutée à des cellules, ou fragments de cellules, exprimant l'ADN codant pour le récepteur marqué par la protéine fluorescente et pour la protéine G marquée par le marqueur, caractérisé en ce que :

- une molécule non fluorescente biologiquement active et agoniste déclenche une transduction de signal détectée par variation de transfert d'énergie entre le récepteur marqué par la protéine fluorescente et la protéine G marquée par le marqueur ;

- une molécule non fluorescente biologiquement active antagoniste inhibe la transduction de signal provoquée par un agoniste et détectée par variation de transfert d'énergie de fluorescence entre le récepteur marqué par la protéine fluorescente et la protéine G marquée par le marqueur.

5

29. Cellules ou fragment de cellules contenant une séquence d'ADN comprenant le gène codant pour une protéine fluorescente fusionnée avec le gène d'une protéine cible, la protéine fluorescente étant choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est supérieur à environ 10 14000M⁻¹cm⁻¹et le rendement quantique de fluorescence est supérieur à environ 0,38, la fusion entre le gène de la protéine fluorescente et le gène de la susdite protéine cible étant telle que

15

- * les propriétés de la protéine cible ne sont pas modifiées par la présence de la protéine fluorescente, à savoir
- * l'interaction entre la protéine cible et le ligand n'est pas modifiée,
- * la fonction de transduction de la réponse n'est pas modifiée,

sous réserve que :

20

- * lorsque la protéine cible est le récepteur des glucocorticoïdes de rat fusionné en N-terminal avec successivement une séquence de purification comportant 6 histidines, un épitope hémagglutinine et une protéine fluorescente et est exprimée dans la lignée cellulaire 1471.1, la protéine fluorescente est différente de la GFP (768 paires de bases du plasmide TU65 avec la mutation S65T),

25

- * lorsque la protéine cible est le récepteur humain des glucocorticoïdes tronqué de ses 131 premiers amino acides, fusionné en C-terminal d'une protéine fluorescente dans les sites Sal I et BamH I et est exprimée dans les cellules Cos-1, ladite protéine fluorescente est différente de la GFP telle que décrite dans l'article de Inouye S. et Tsuji, F. I., 1994, Febs Letters, Vol. 341, p.277-280,

30

- * lorsque la protéine cible est la sous-unité NMDA R1 de rat exprimé dans les cellules HEK 293 fusionnée en C-terminal avec une protéine fluorescente, la protéine fluorescente est différente de celle constituée par les acides aminés 2-238 de la GFP sauvage,

35

- * lorsque la protéine cible est un récepteur ou un fragment de récepteur de messagers secondaires intracellulaires, la protéine fluorescente est différente de la GFP et de ses dérivés.

5 30. Cellules ou fragment de cellules contenant une séquence d'ADN comprenant le gène codant pour une substance fluorescente fusionnée avec le gène d'un ligand, la substance fluorescente étant choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est supérieur à environ $14000\text{M}^{-1}\text{cm}^{-1}$ et le rendement quantique de fluorescence est supérieur à environ 0,38, la fusion entre le gène de la substance fluorescente et le gène du ligand étant tel que

- 10
 - * les propriétés du ligand ne sont pas modifiées par la présence de la substance fluorescente, à savoir
 - * l'interaction entre la protéine cible et le ligand n'est pas modifiée,
 - * la fonction de transduction de la réponse n'est pas modifiée.

15 31. Cellules ou fragment de cellules contenant une séquence d'ADN comprenant

20

- le gène codant pour un ligand constitué par la protéine G fusionnée avec une substance fluorescente choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est supérieur à environ $14000\text{M}^{-1}\text{cm}^{-1}$ et le rendement quantique de fluorescence est supérieur à environ 0,38,

25

- et éventuellement le gène codant pour une protéine fluorescente fusionnée avec le gène d'une protéine cible constituée par un récepteur, la protéine fluorescente étant choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est supérieur à environ $14000\text{M}^{-1}\text{cm}^{-1}$ et le rendement quantique de fluorescence est supérieur à environ 0,38,

30

- la fusion entre le gène de la substance fluorescente et le gène de la susdite protéine G, et éventuellement la fusion entre le gène de la protéine fluorescente et le gène du récepteur étant telle que

- 35
 - * les propriétés de la protéine G ne sont pas modifiées par la présence de la substance fluorescente,
 - * les propriétés du récepteur ne sont pas modifiées par la présence de la protéine fluorescente, à savoir
 - * l'interaction entre la protéine G et le récepteur n'est pas modifiée,

- * la fonction de transduction de la réponse n'est pas modifiée, sous réserve que :
- * lorsque la protéine cible est le récepteur des glucocorticoïdes de rat fusionné en N-terminal avec successivement une séquence de purification comportant 6 histidines, un épitope hémagglutinine et une protéine fluorescente et est exprimée dans la lignée cellulaire 1471.1, la protéine fluorescente est différente de la GFP (768 paires de bases du plasmide TU65 avec la mutation S65T),
- * lorsque la protéine cible est le récepteur humain des glucocorticoïdes tronqué de ses 131 premiers amino acides, fusionné en C-terminal d'une protéine fluorescente dans les sites Sal I et BamH I et est exprimée dans les cellules Cos-1, ladite protéine fluorescente est différente de la GFP telle que décrite dans l'article de Inouye S. et Tsuji, F. I., 1994, Febs Letters, Vol. 341, p.277-280,
- * lorsque la protéine cible est la sous-unité NMDA R1 de rat exprimé dans les cellules HEK 293 fusionnée en C-terminal avec une protéine fluorescente, la protéine fluorescente est différente de celle constituée par les acides aminés 2-238 de la GFP sauvage,
- * lorsque la protéine cible est un récepteur ou un fragment de récepteur de messagers secondaires intracellulaires, la protéine fluorescente est différente de la GFP et de ses dérivés.

32. Trousse ou nécessaire pour la détection et la quantification d'interactions non covalentes entre une protéine cible marquée par une protéine fluorescente et l'un de ses ligands marqué par un marqueur constitué :

- soit par une molécule susceptible d'absorber la lumière émise par la protéine fluorescente,
- soit par une substance fluorescente,

laquelle protéine fluorescente est choisie parmi les protéines fluorescentes issues ou dérivées de protéines autofluorescentes de cnidaires, dont le coefficient d'extinction moléculaire est supérieur à environ $14000M^{-1}cm^{-1}$ et le rendement quantique de fluorescence est supérieur à environ 0,38, notamment choisie parmi :

- la protéine fluorescente verte (GFP), ou

- des variants dérivés de la GFP, par addition, délétion ou substitution d'un ou plusieurs acides aminés, sous réserve que ces variants conservent la propriété de fluorescence,

5 - ou des fragments de la GFP, ou de fragments des susdits variants, sous réserve que ces fragments conservent la propriété de fluorescence et son ligand marqué par une substance fluorescente, ladite trousse comprenant :

10 - la protéine cible fusionnée avec une protéine fluorescente ou une lignée cellulaire stable susceptible d'exprimer la protéine fusionnée avec une protéine fluorescente ou un plasmide contenant la séquence nucléique codant pour ladite protéine cible fusionnée avec une protéine fluorescente telle que définie ci-dessus,

15 - le ligand marqué par le susdit marqueur,
- les tampons et milieux nécessaires au transfert d'énergie entre la susdite protéine et le susdit ligand.

15

33. Trousse ou nécessaire pour la détection et la quantification d'interactions non covalentes entre une protéine cible marquée par une protéine fluorescente (n° 1) et l'un de ses ligand marqué par une substance fluorescente correspondant à une protéine fluorescente (n° 2), la protéine fluorescente (n° 1) étant choisie parmi la protéine fluorescente EYFP ou EGFP et le ligand étant marqué par la protéine fluorescente (n° 2) ECFP ou la protéine fluorescente (n° 1) étant ECFP et le ligand étant marqué par la protéine fluorescente (n° 2) EYFP ou EGFP, ladite trousse comprenant :

20

- soit un plasmide contenant un acide nucléique codant pour la protéine cible fusionnée avec une protéine fluorescente (n° 1), et

- * un plasmide contenant un acide nucléique codant pour le ligand fusionné avec une protéine fluorescente (n° 2), ou
- * un ligand fusionné avec une protéine fluorescente (n° 2), obtenu par voie recombinante et purifié,

25

- soit une lignée cellulaire stable susceptible d'exprimer la protéine cible fusionnée avec une protéine fluorescente (n° 1), et

- * une lignée cellulaire stable susceptible d'exprimer le ligand fusionné avec une protéine fluorescente (n° 2), ou
- * un ligand fusionné avec une protéine fluorescente (n° 2), obtenu par voie recombinante et purifié,

30

- les tampons et milieux nécessaires au transfert d'énergie entre la susdite protéine et le susdit ligand.

34. Trousse ou nécessaire pour la détection et la quantification d'interactions non covalentes entre une protéine cible constituée par un récepteur couplé à la protéine G marquée par une protéine fluorescente (n° 1) et la protéine G marquée par une substance fluorescente correspondant à une protéine fluorescente (n° 2), la protéine fluorescente (n° 1) étant choisie parmi la protéine fluorescente EYFP ou EGFP et la protéine G étant marquée par la protéine fluorescente (n° 2) ECFP ou la protéine fluorescente (n° 1) étant ECFP et la protéine G étant marquée par la protéine fluorescente (n° 2) EYFP ou EGFP, ladite trousse comprenant :

- soit un plasmide contenant un acide nucléique codant pour le récepteur fusionné avec une protéine fluorescente (n° 1), et
 - * un plasmide contenant un acide nucléique codant pour la protéine G fusionnée avec une protéine fluorescente (n° 2), ou
 - * la protéine G fusionnée avec une protéine fluorescente (n° 2), obtenue par voie recombinante et purifiée,
- soit une lignée cellulaire stable susceptible d'exprimer le récepteur fusionné avec une protéine fluorescente (n° 1), et
 - * une lignée cellulaire stable susceptible d'exprimer la protéine G fusionnée avec une protéine fluorescente (n° 2), ou
 - * la protéine G fusionnée avec une protéine fluorescente (n° 2), obtenue par voie recombinante et purifiée,
- les tampons et milieux nécessaires au transfert d'énergie entre le susdit récepteur et la susdite protéine G.

atg gtg agg aag ggc gag gag ctc acc ggg gtg ccc atc ctg gtc gag ctg gac ggc gag gta
M V S K Q E L F T G V P I L V E L D G D V

aac ggc cac aag ttc agg gtg tcc ggc gag ggc gag ggc gat gcc acc tac ggc aag ctg acc ctg aag
N G H K F S V S G E G D A T Y G K L T L K

ttc atc tgc acc acc ggc aag ctc agg ccc gtg ccc tgg ccc acc ctc gtg acc acc ctg acc tac ggc gtg
F I C T T G K L P V P W P T L V T T L T Y G V

cag tgc ttc agc cgc tac ccc gac cac atg aag cag cac gac ttc aag tcc gcc atg ccc gaa ggc
Q C F S R Y P D H M K Q H D F F K S A M P E G

tac gtc cag gag cgc acc atc ttc aag gag gac ggc aac tac aag acc cgc ggc gag gtg aag ttc
Y V Q E R T I F F K D D G N Y K T R A E V K F

gag ggc gac acc ctg gtg aac cgc atc gag ctg aag ggc atc gac ttc aag gag gac ggc aac atc ctg
E G D T L V N R I E L K G I D F K E D G N I L

ggg cac aag ctg gag tac aac tac aac agc cac gtc tat atc atg gcc gac aag cag aag aac ggc
G H K L E Y N S H N V Y I M A D K Q K N G

atc aag gtg aac ttc aag atc cgc cac aac atc gag gag gac ggc agg gtc gag ctc gac cac tac gag
I K V N F K I R H N I E D G S V Q L A D H Y Q

cag aac acc ccc atc ggc gac ggc ccc gtg ctg ccc gac aac cac tac ctg agc acc cag tcc gcc
Q N T P I G D G P V L L P D N H Y L S T Q S A

ctg agc aaa gag ccc aac gag aag cgc gat cac atg gtc ctg ctg gag ttc gtg acc gcc ggg atc
L S K D P N E K R D H M V L E F V T A A G I

act ctc ggc atg gag tac tca gag tac tca gag ctc aag ctt cga att ctg gag tcg acg
T L G M D E L Y K S D L E L K R I L Q S T

gta ccg cgg gcc cggtt gat cca ccc gat cta gat aac tga
V P R A R D P D L D N * *

Figure 1

09/445205

WO 98/55873

PCT/FR98/01136

2/22

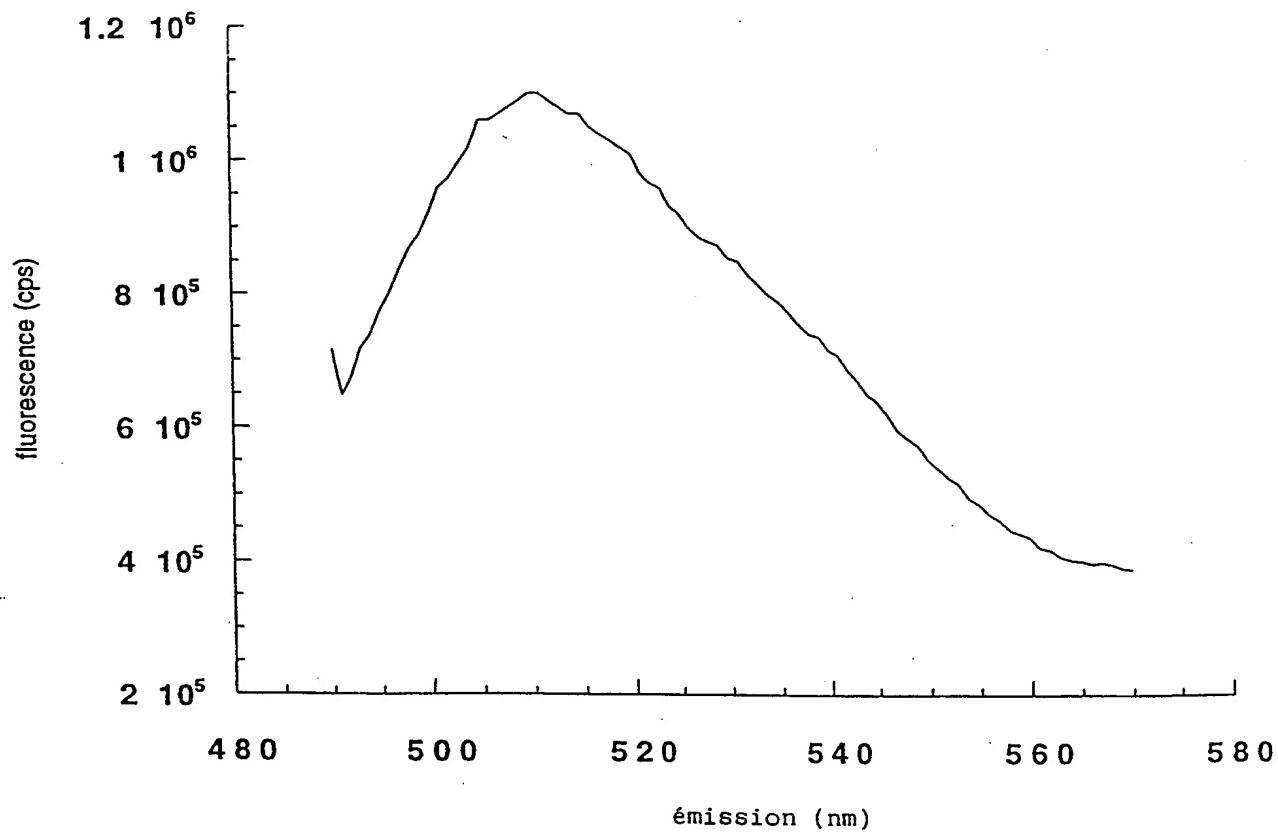


Figure 2

09/445205

WO 98/55873

PCT/FR98/01136

3/22

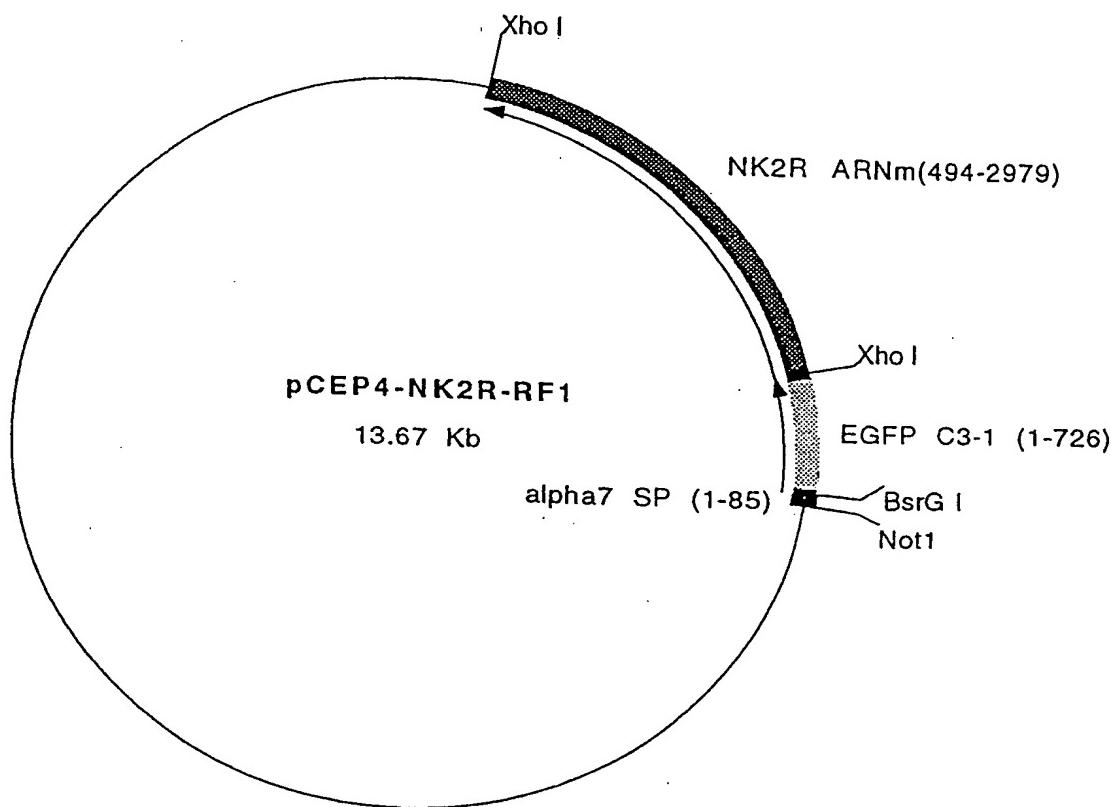


Figure 3

4/22

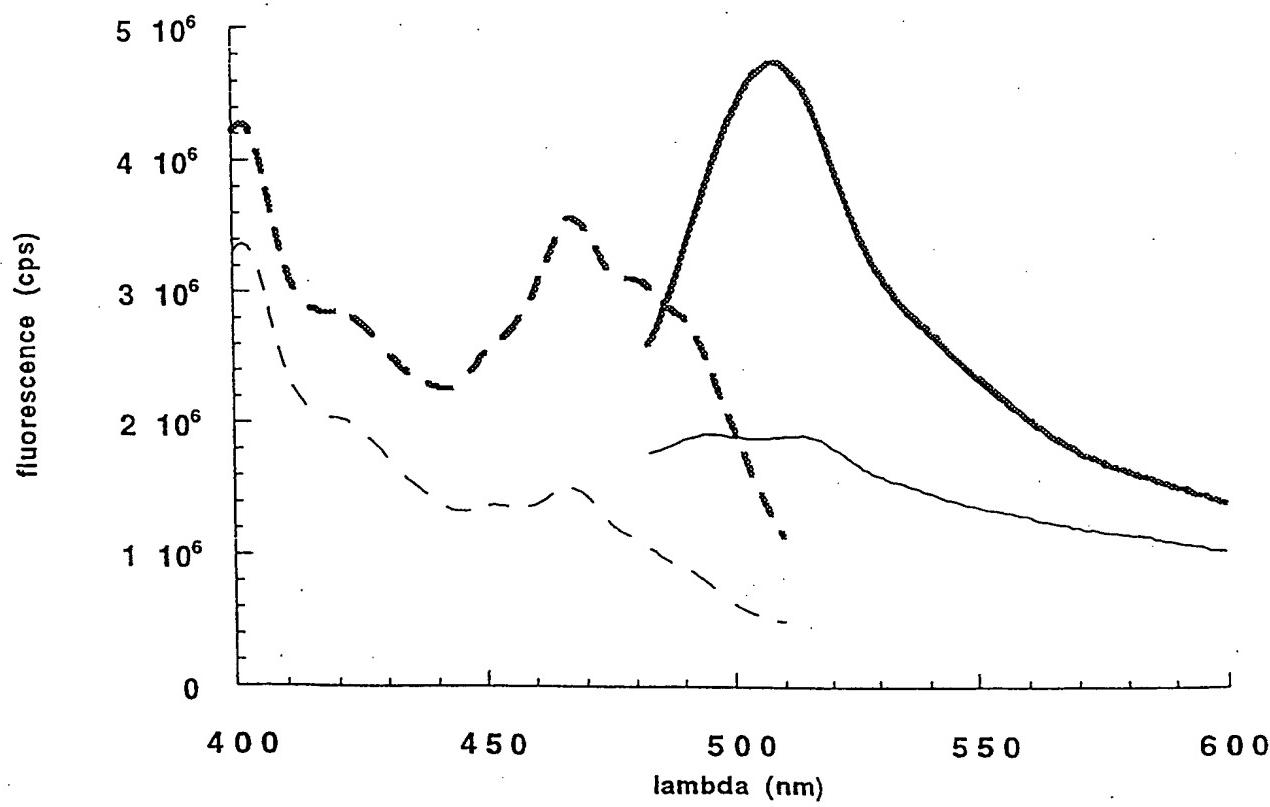


Figure 4a

5/22

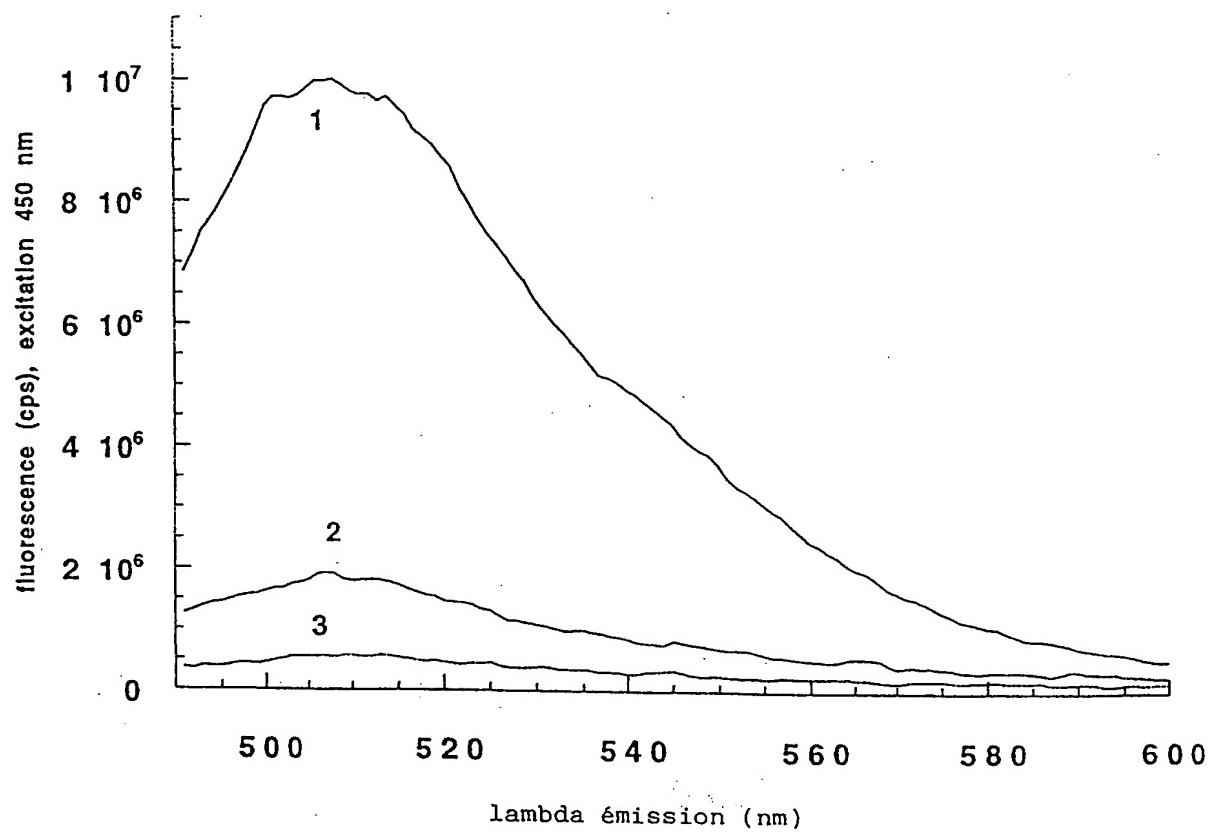


Figure 4b

09/445205

WO 98/55873

PCT/FR98/01136

6/22

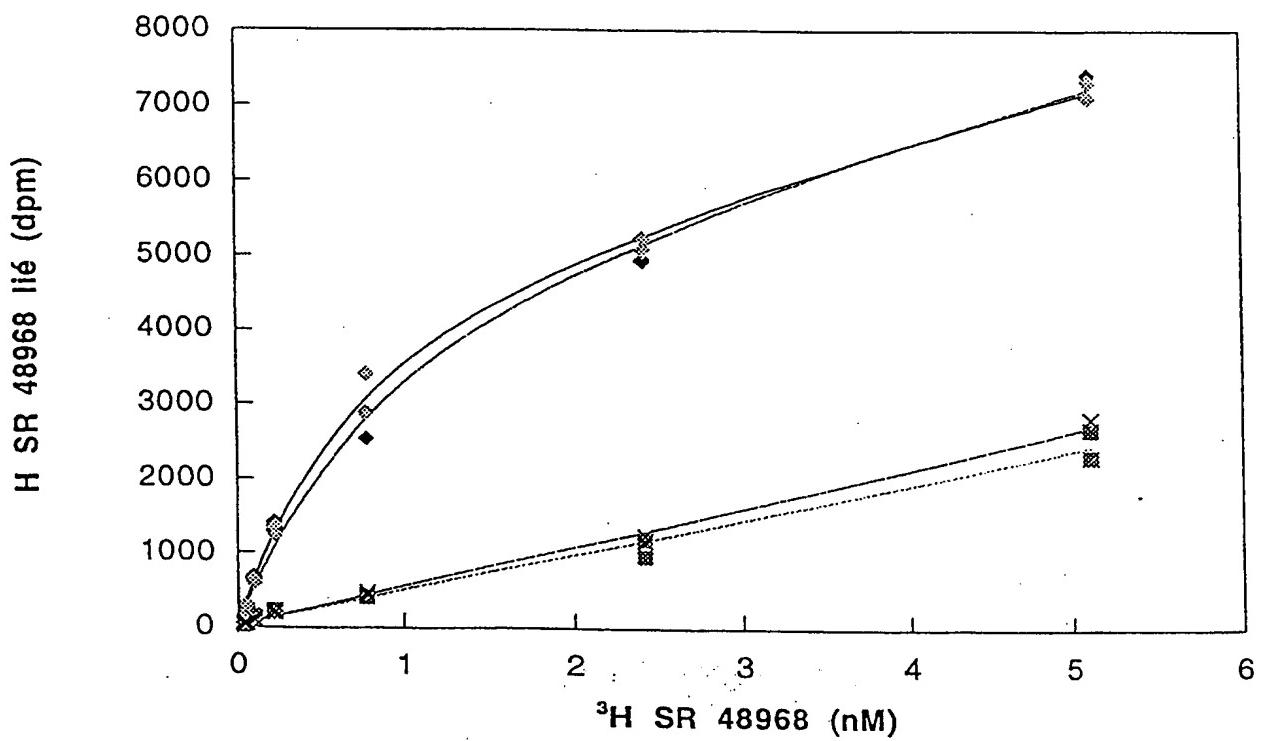


Figure 5a

7/22

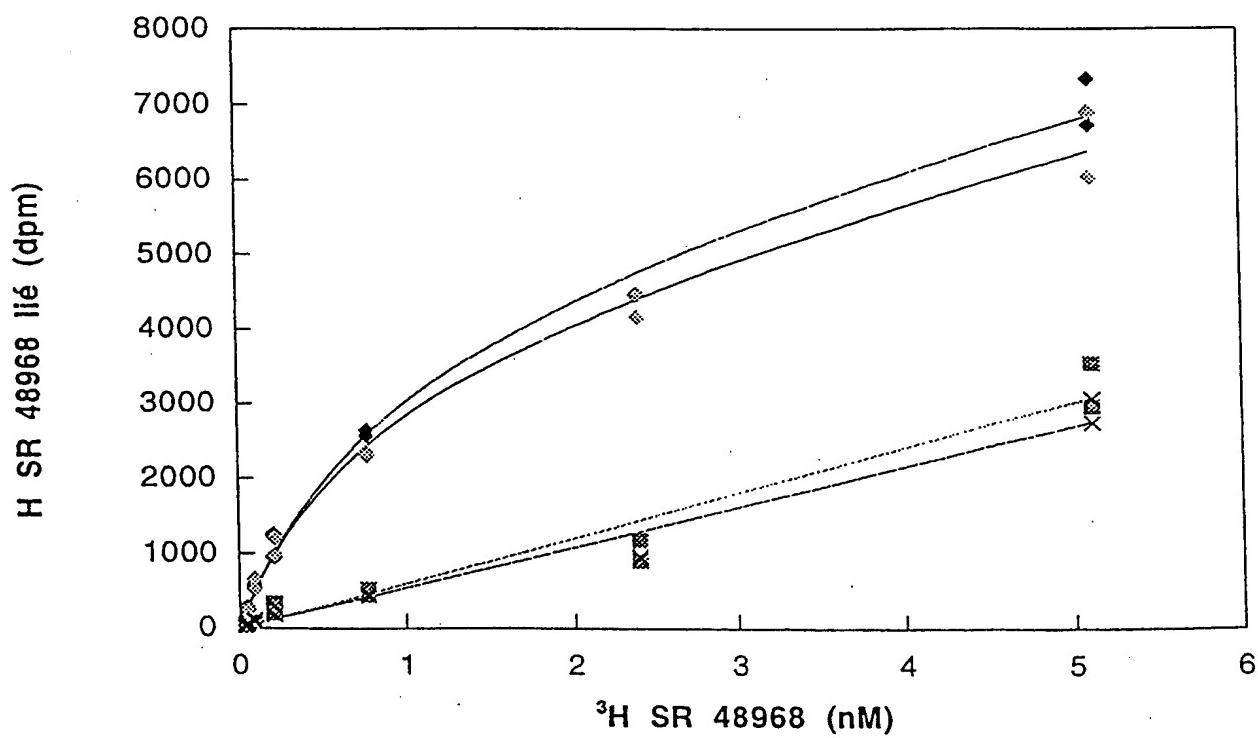


Figure 5b

09/445205

WO 98/55873

PCT/FR98/01136

8/22

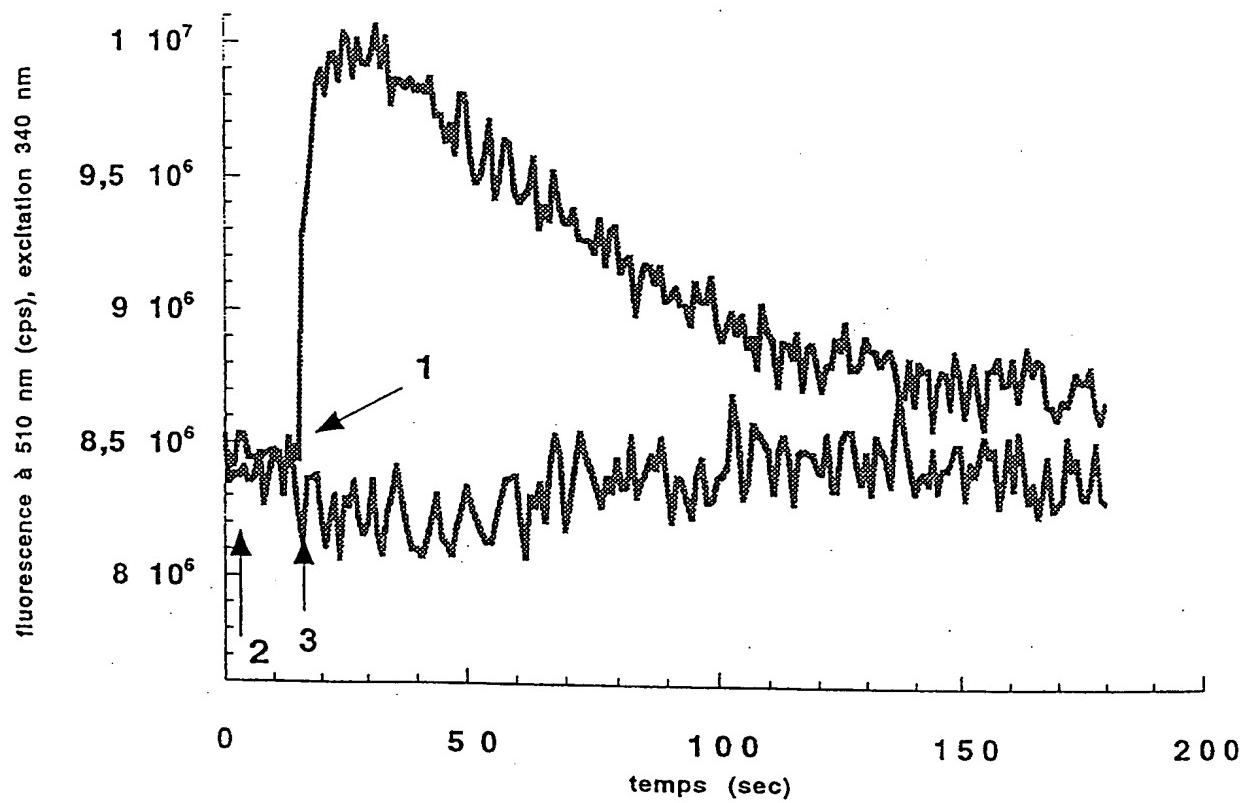


Figure 6a

09/445205

WO 98/55873

PCT/FR98/01136

9/22

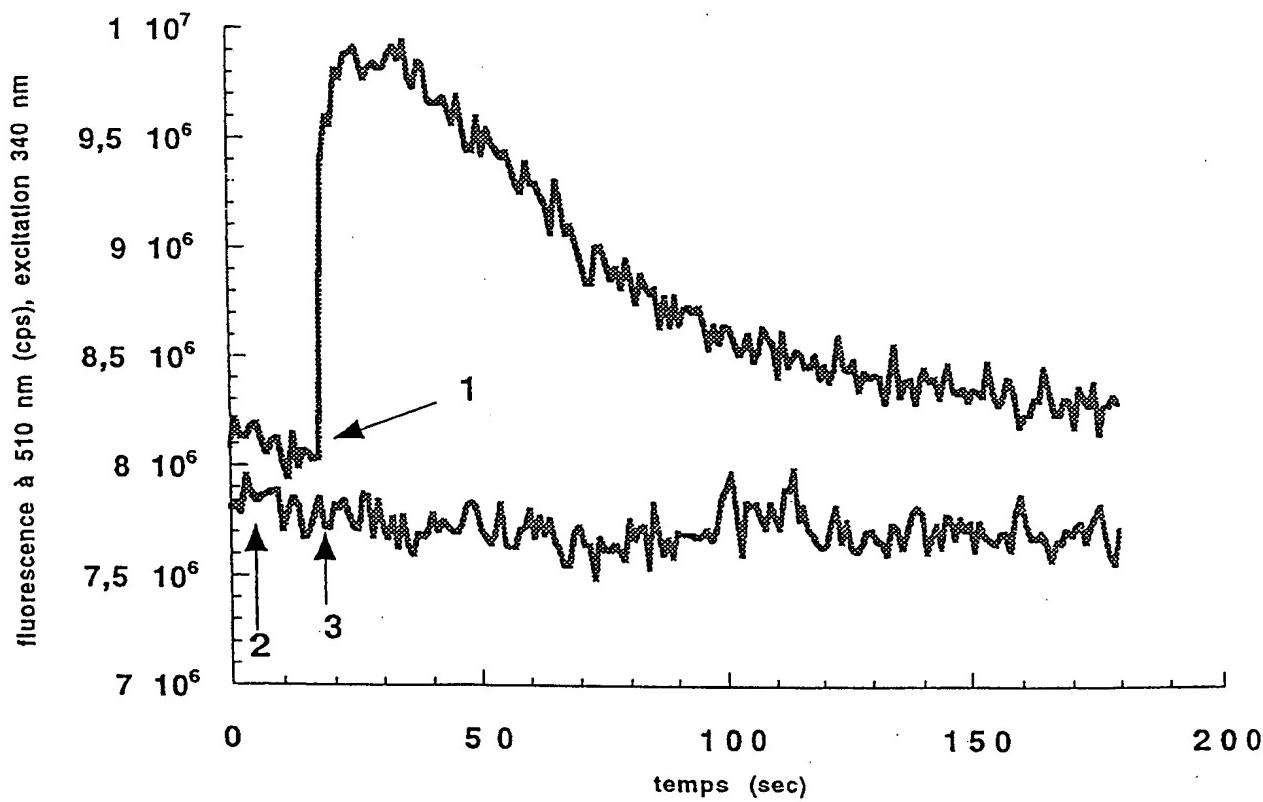


Figure 6b

10/22

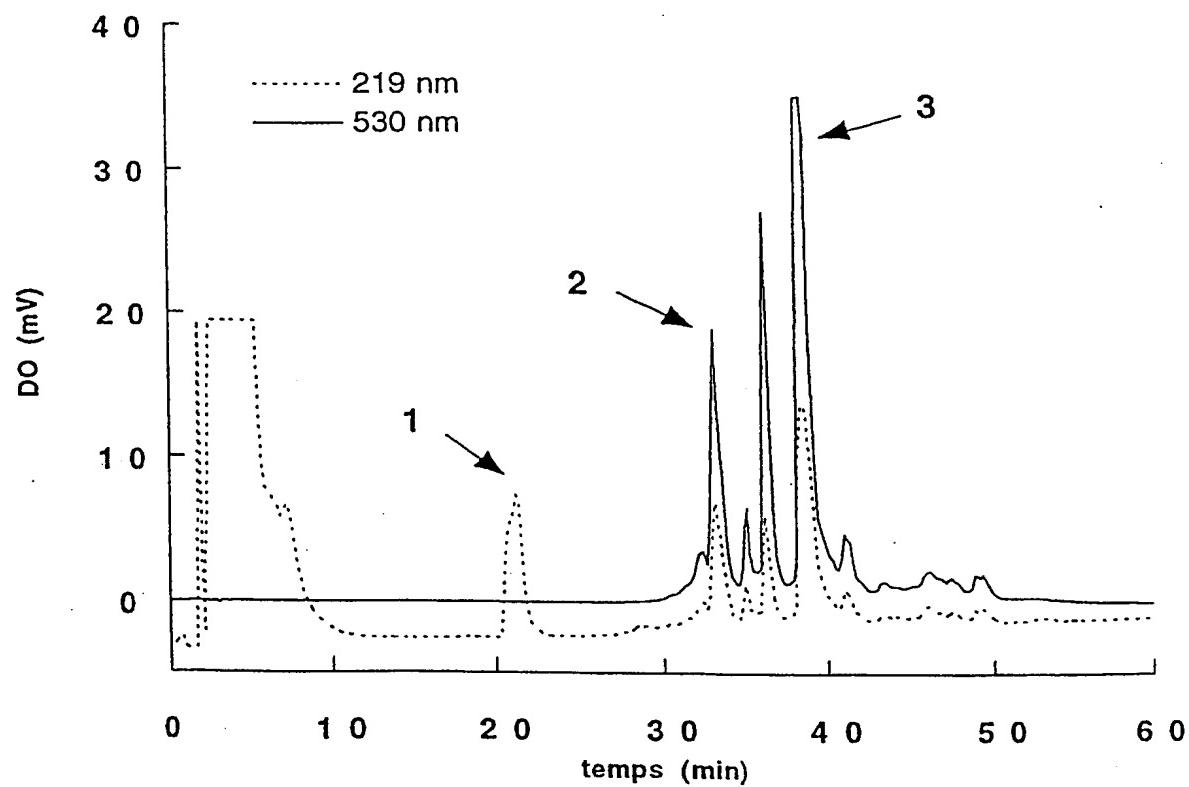


Figure 7

09/445205

WO 98/55873

PCT/FR98/01136

11/22

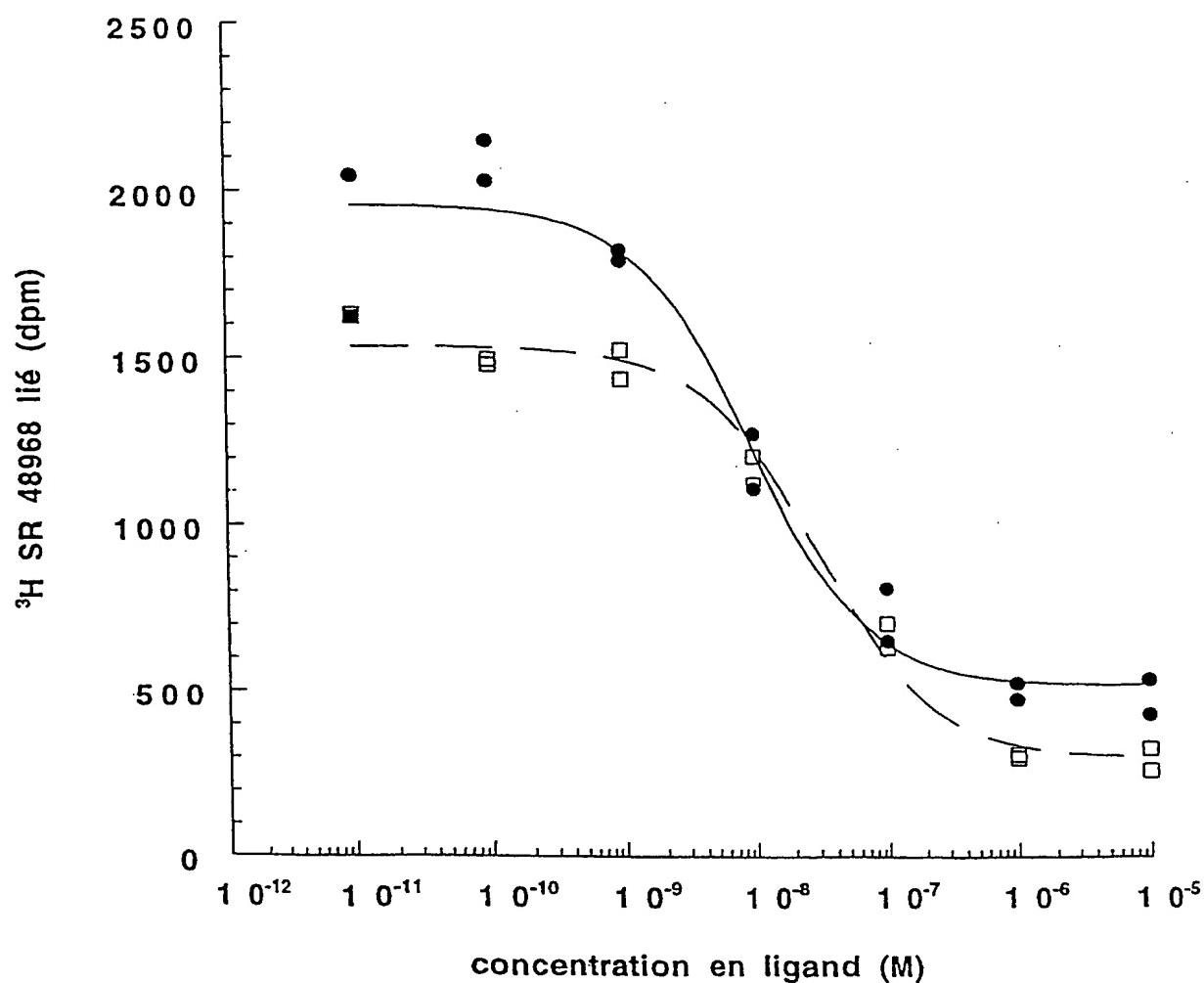


Figure 8a

09/445205

WO 98/55873

PCT/FR98/01136

12/22

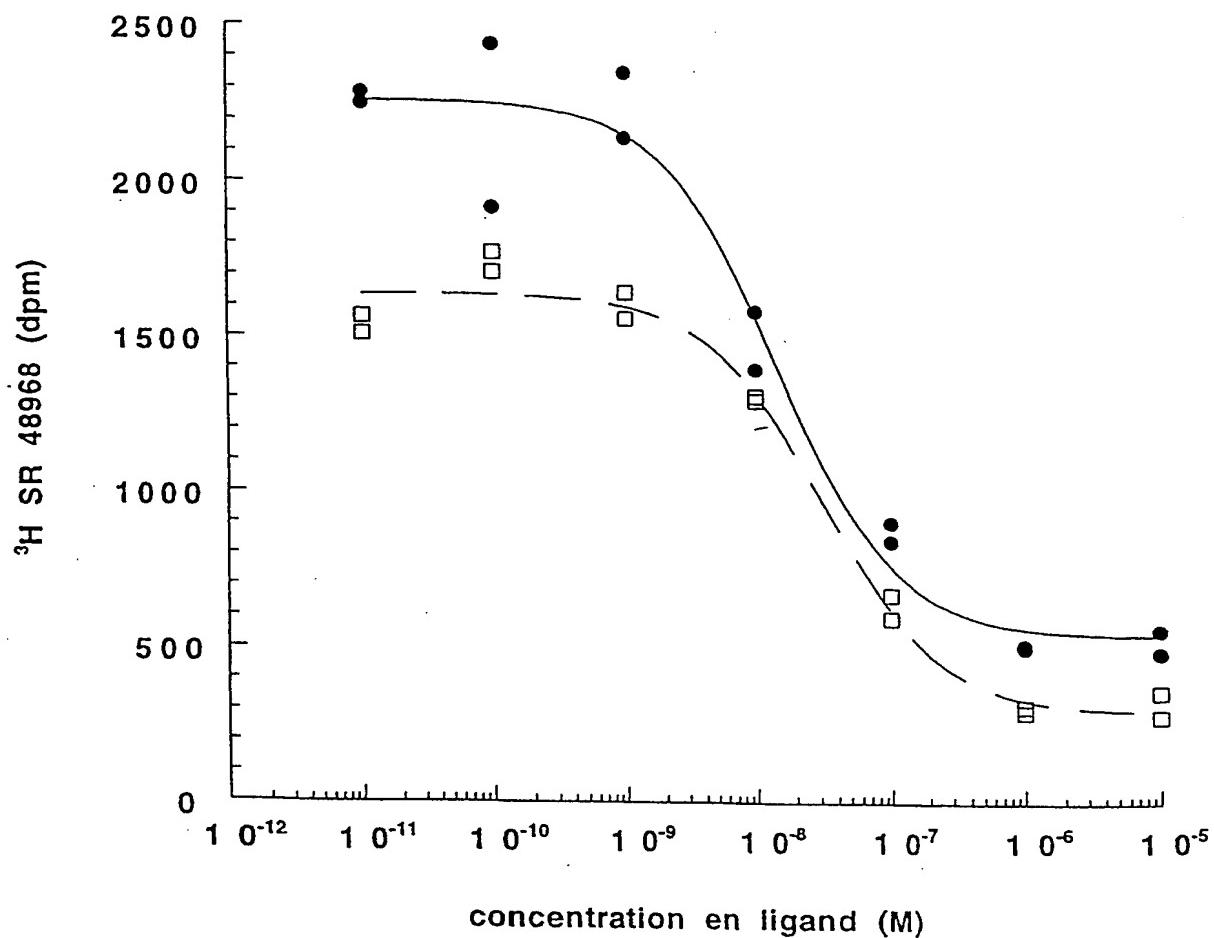


Figure 8b

09/445205

WO 98/55873

PCT/FR98/01136

13/22

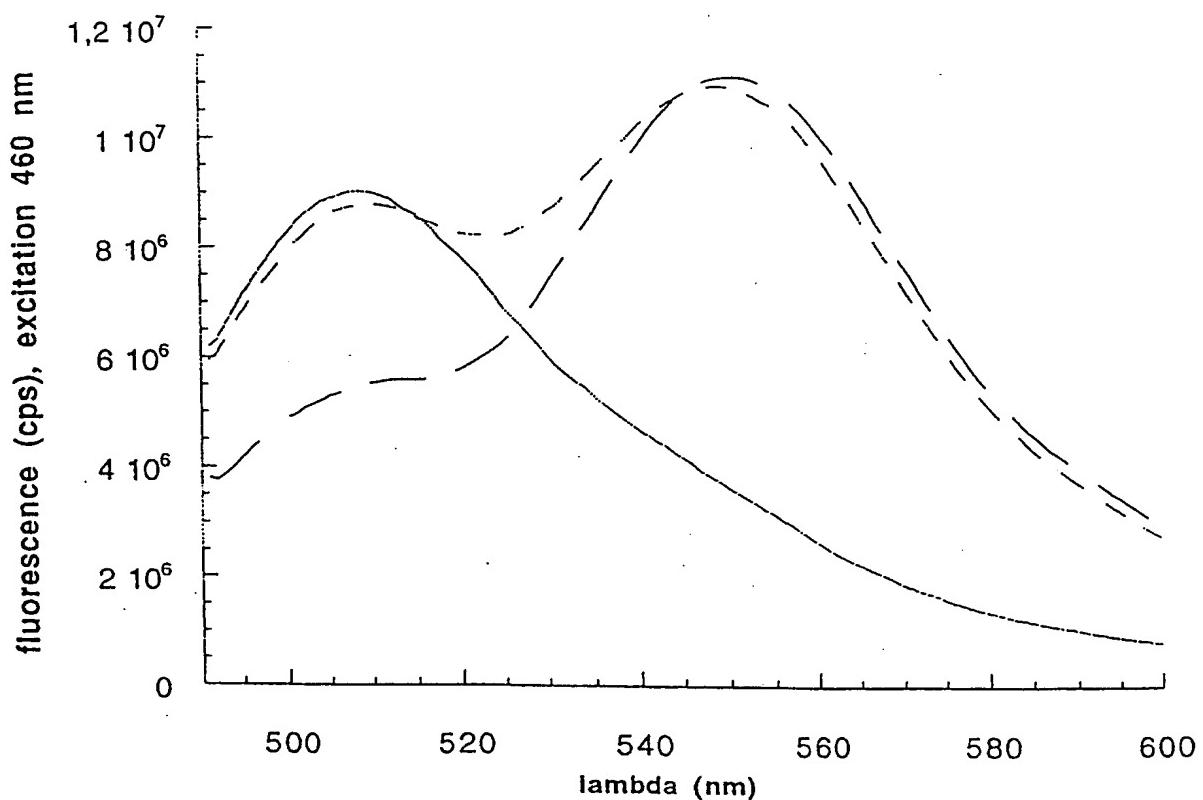


Figure 9

09/445205

WO 98/55873

PCT/FR98/01136

14/22

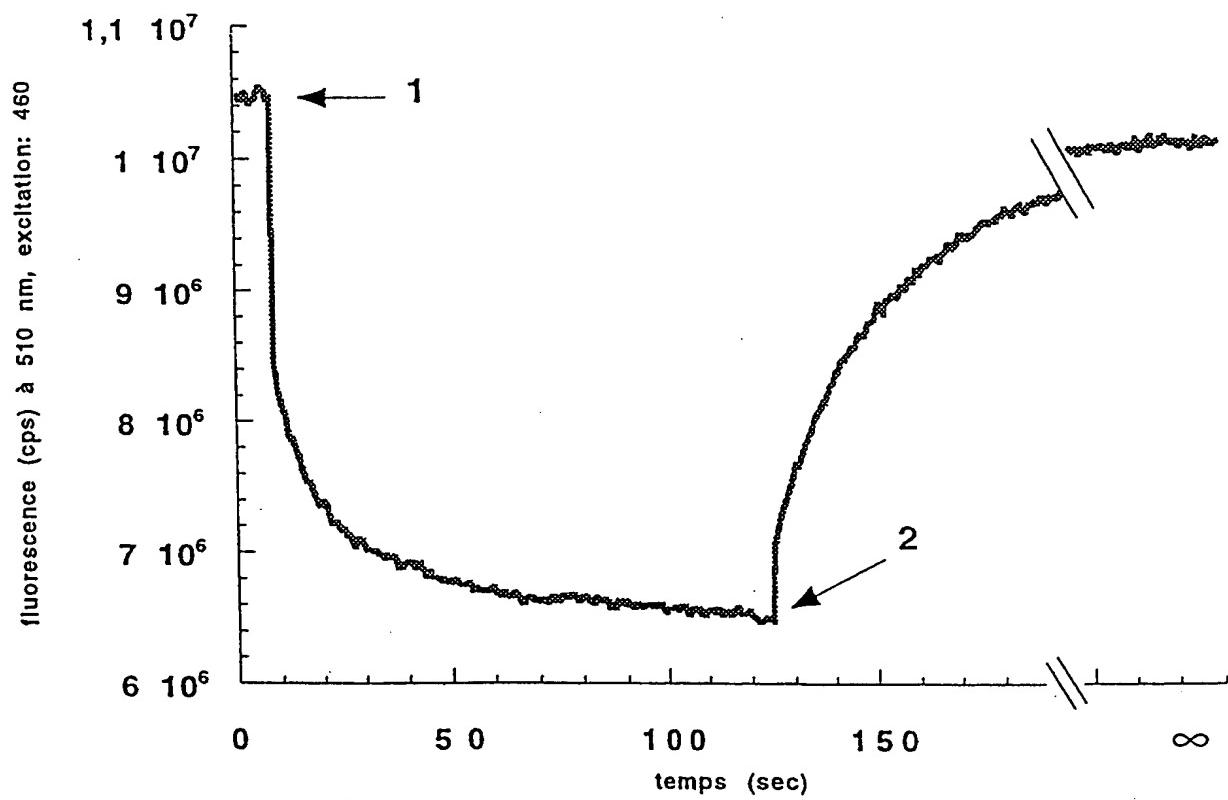


Figure 10

09/445205

WO 98/55873

PCT/FR98/01136

15/22

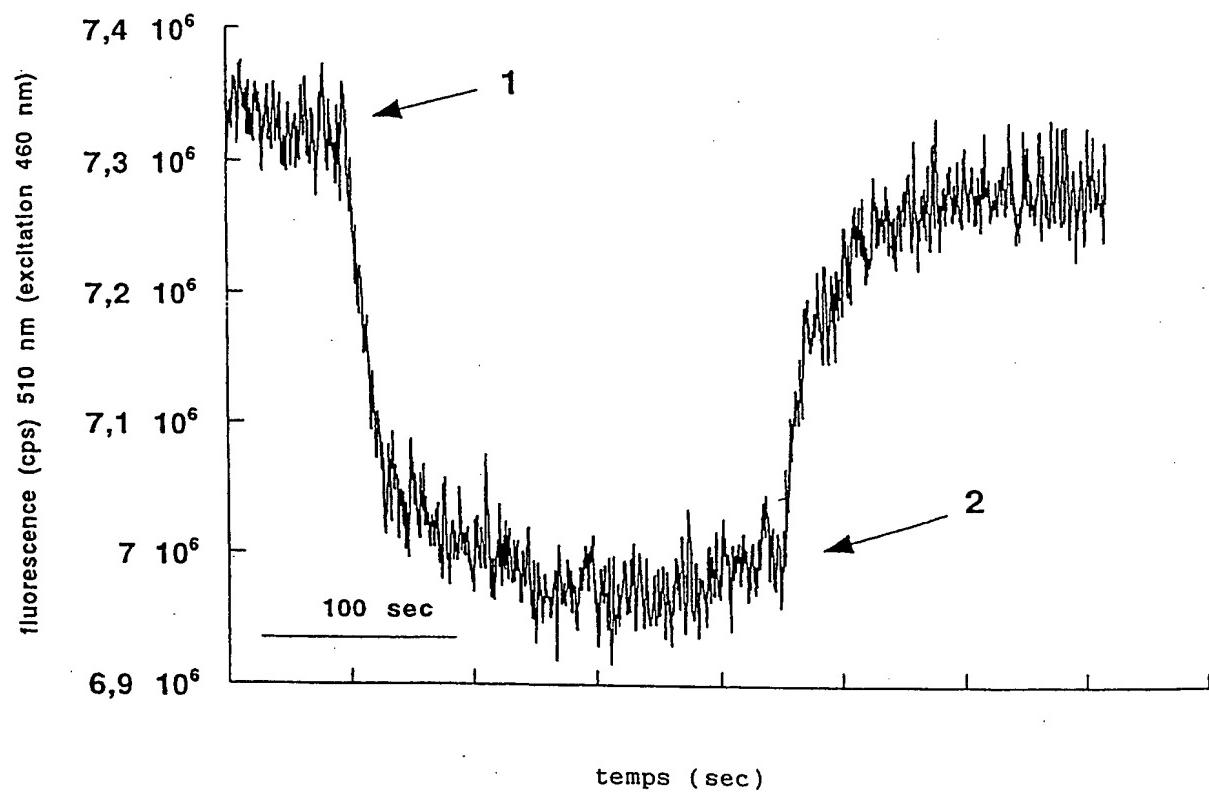


Figure 11a

09/445205

WO 98/55873

PCT/FR98/01136

16/22

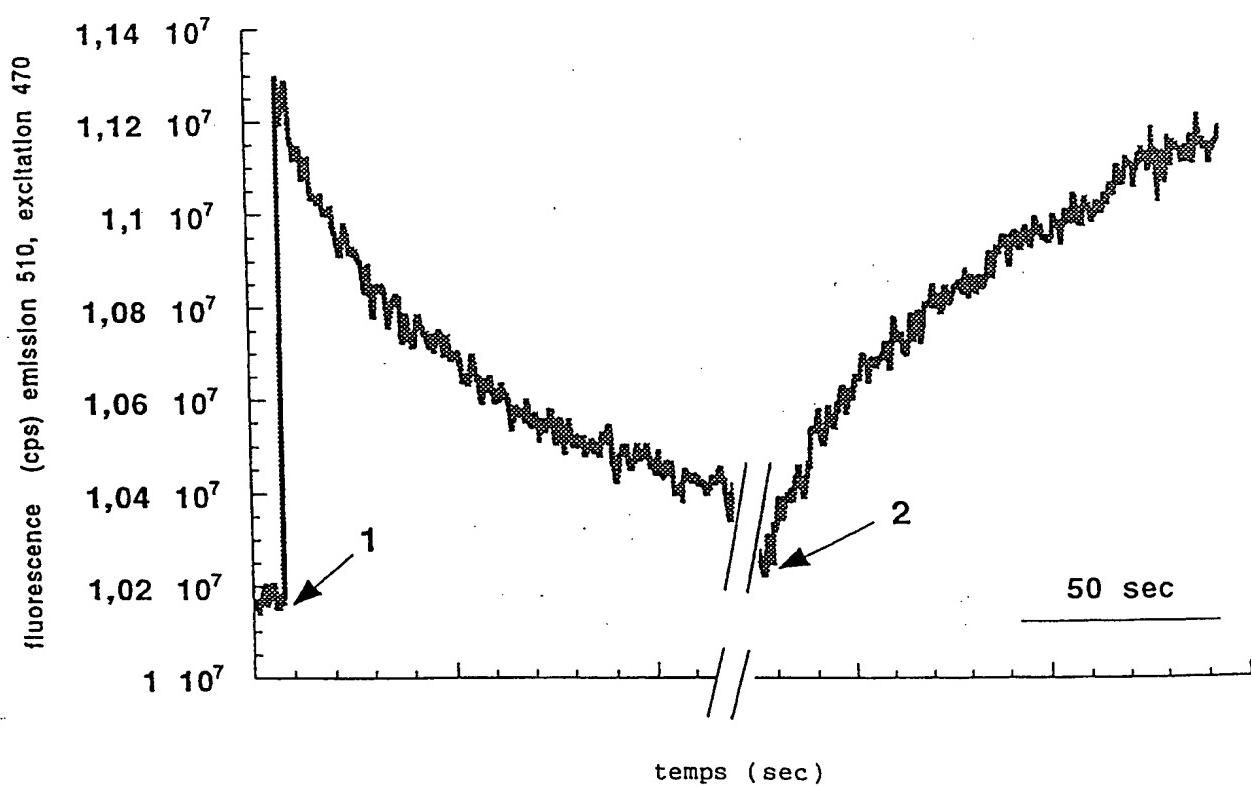


Figure 11b

17/22

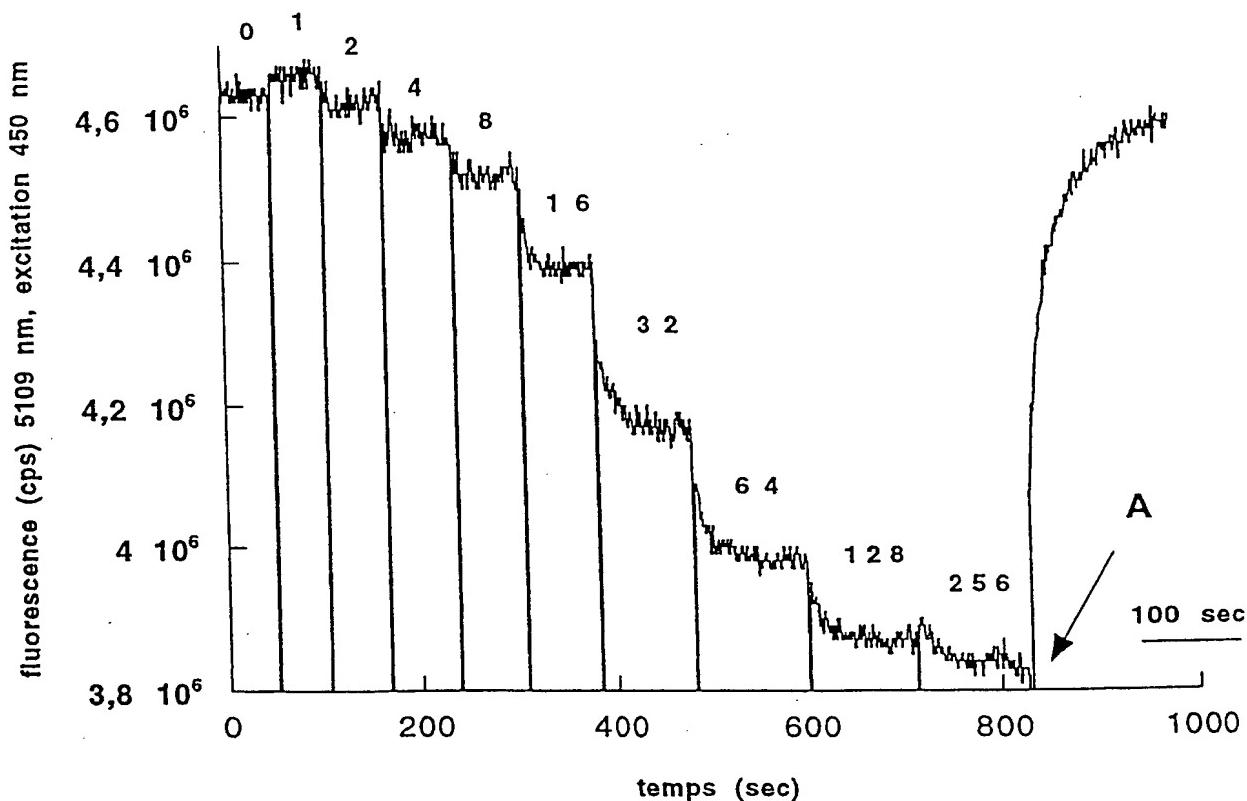


Figure 12

18/22

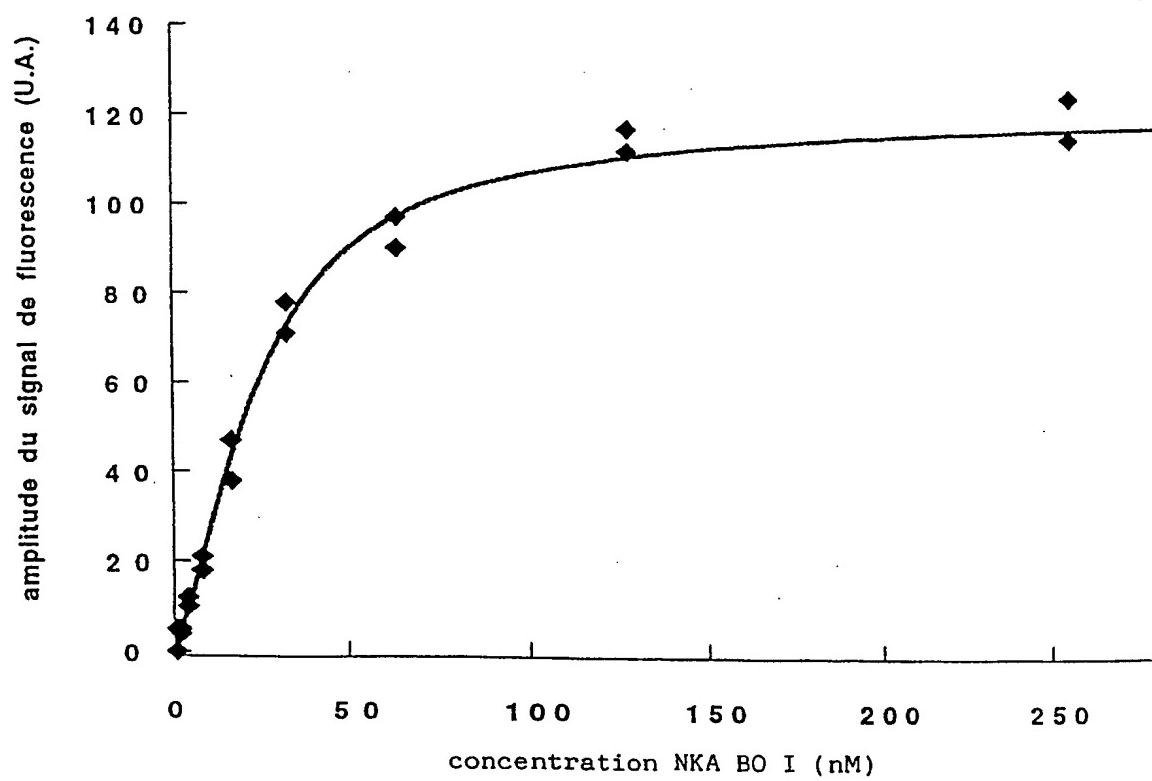
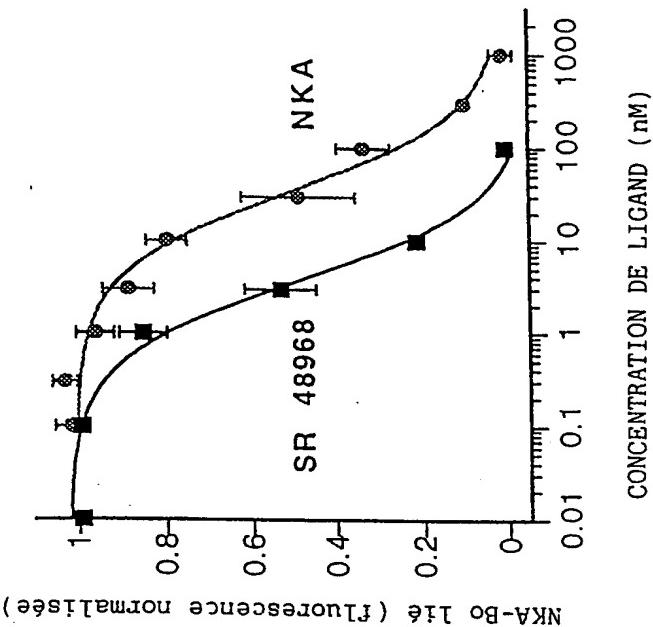
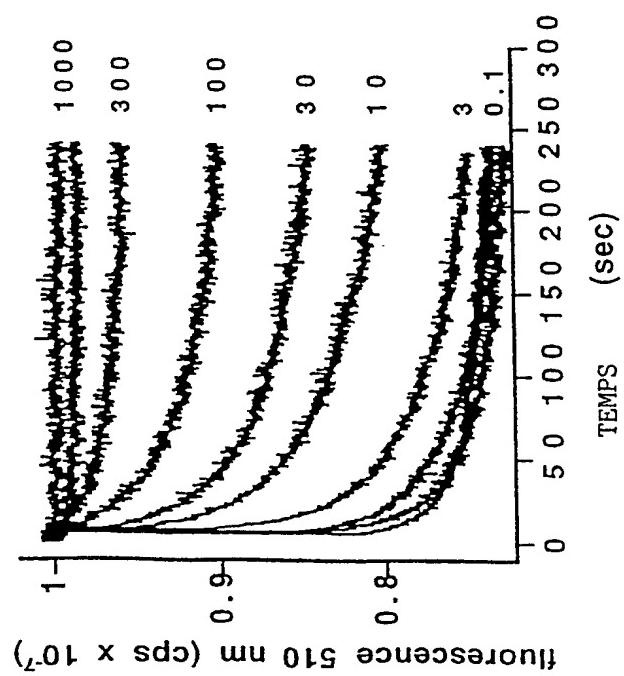


Figure 13

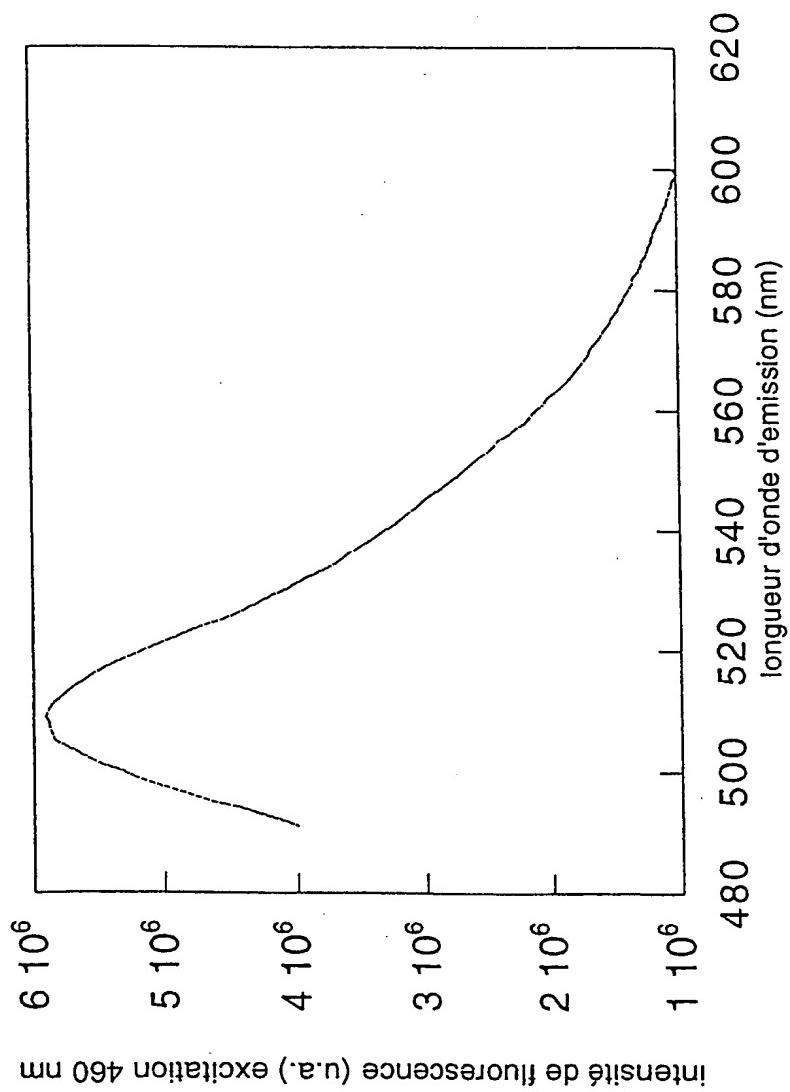
19/22

Figure 14
Figure 15



20/22

Figure 16



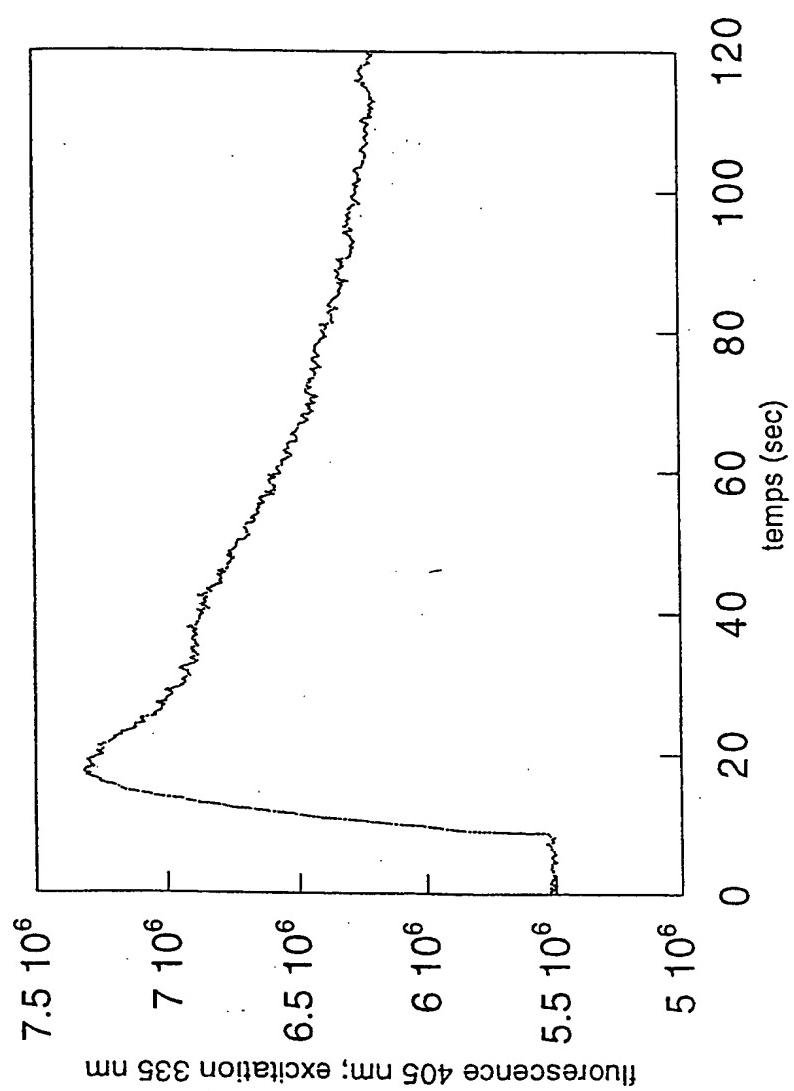
09/445205

WO 98/55873

PCT/FR98/01136

21/22

Figure 17



09/445205

WO 98/55873

PCT/FR98/01136

22/22



Figure 18

RAPPORT DE RECHERCHE INTERNATIONALE

Demande Internationale No

PCT/FR 98/01136

A. CLASSEMENT DE L'OBJET DE LA DEMANDE
 CIB 6 G01N33/58 C12Q1/25 C12N5/10 C12Q1/37

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)
 CIB 6 G01N C12Q C12N

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés)

C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie *	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
P, X	WO 97 28261 A (UNIV. CALIFORNIA; AURORA BIOSCIENCES CORP) 7 août 1997 voir le document en entier ---	1-34
P, X	MIYAWAKI A ET AL: "Fluorescent indicators for Ca-2+ based on green fluorescent proteins and calmodulin." NATURE (LONDON) 388 (6645). 1997. 882-887. ISSN: 0028-0836, XP002058386 voir page 886, colonne 2 ---	1-34 -/-

Voir la suite du cadre C pour la fin de la liste des documents

Les documents de familles de brevets sont indiqués en annexe

* Catégories spéciales de documents cités:

- "A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent
- "E" document antérieur, mais publié à la date de dépôt international ou après cette date
- "L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)
- "O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens
- "P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

"T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention

"X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément

"Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier

"&" document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

30 décembre 1998

Date d'expédition du présent rapport de recherche internationale

08/01/1999

Nom et adresse postale de l'administration chargée de la recherche internationale
 Office Européen des Brevets, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
 Fax: (+31-70) 340-3016

Fonctionnaire autorisé

Fletcher, A

RAPPORT DE RECHERCHE INTERNATIONALE

Demande Internationale No

PCT/FR 98/01136

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
Y	ROMOSER V A ET AL: "Detection in living cells of Ca-2+-dependent changes in the fluorescence emission of an indicator composed of two green fluorescent protein variants linked by a calmodulin - binding sequence: A new class of fluorescent indicators." JOURNAL OF BIOLOGICAL CHEMISTRY 272 (20). 1997. 13270-13274. ISSN: 0021-9258, XP002058387 voir le document en entier ---	1-34
Y	PETUSHKOV V N ET AL: "Direct measurement of excitation transfer in the protein complex of bacterial luciferase hydroxyflavin and the associated yellow fluorescence proteins from Vibrio fischeri Y1." BIOCHEMISTRY 35 (25). 1996. 8413-8418. ISSN: 0006-2960, XP002058388 voir le document en entier ---	1-34
Y	MITRA R D ET AL: "Fluorescence resonance energy transfer between blue-emitting and red-shifted excitation derivatives of the green fluorescent protein." GENE (AMSTERDAM) 173 (1). 1996. 13-17. ISSN: 0378-1119, XP004042847 voir le document en entier ---	1-34
X	WO 91 01305 A (UNIV WALES MEDICINE) 7 février 1991	1-4, 6, 10, 11, 17, 18, 32
Y	voir exemples 3, 7	5, 7-9, 13-16, 19-21, 29, 33
Y	---	1-11, 13-21, 29, 32, 33
Y	WO 96 41166 A (UNIV CALIFORNIA ; TSIEN ROGER Y (US); GONZALEZ JESUS E III (US)) 19 décembre 1996 voir page 33, ligne 30 - ligne 35 ---	1-11, 13-21, 29, 32, 33
Y	US 5 342 789 A (CHICK WILLIAM L ET AL) 30 août 1994 voir le document en entier ---	1-11, 13-21, 29, 32, 33
Y	EP 0 552 108 A (UNIV MARYLAND) 21 juillet 1993 voir le document en entier. ---	1-11, 13-21, 29, 32, 33
		-/-

RAPPORT DE RECHERCHE INTERNATIONALE

Demande Internationale No

PCT/FR 98/01136

C.(suite) DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
Y	US 5 439 797 A (TSIEN ROGER Y ET AL) 8 août 1995 voir le document en entier ---	1-11, 13-21, 29,32,33
A	US 5 625 048 A (TSIEN ROGER Y ET AL) 29 avril 1997 voir le document en entier -----	1

RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Demande Internationale No

PCT/FR 98/01136

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)			Date de publication
WO 9728261 A	07-08-1997	AU EP	2250497 A 0877805 A		22-08-1997 18-11-1998
WO 9101305 A	07-02-1991	AU CA EP JP US	6054590 A 2064766 A 0484369 A 5501862 T 5683888 A		22-02-1991 23-01-1991 13-05-1992 08-04-1993 04-11-1997
WO 9641166 A	19-12-1996	US AU CA EP	5661035 A 6264396 A 2223927 A 0834074 A		26-08-1997 30-12-1996 19-12-1996 08-04-1998
US 5342789 A	30-08-1994	EP WO	0505479 A 9109312 A		30-09-1992 27-06-1991
EP 0552108 A	21-07-1993	CA JP US	2087413 A 6066802 A 5631169 A		18-07-1993 11-03-1994 20-05-1997
US 5439797 A	08-08-1995	AT DE EP WO	170980 T 69130171 D 0537270 A 9200388 A		15-09-1998 15-10-1998 21-04-1993 09-01-1992
US 5625048 A	29-04-1997	AU EP JP WO US	4155096 A 0804457 A 10509881 T 9623810 A 5777079 A		21-08-1996 05-11-1997 29-09-1998 08-08-1996 07-07-1998